Three essays about the community mental health in China

Fei Deng

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy.

January 2023

Abstract

People with schizophrenia have long-term impairments and suffer from severe social cognition deficits. This thesis used multiple sources of data (empirical data, extracted data from literatures), applied econometric models, and described the living conditions, treatment situation, and effectiveness of treatment of patients with schizophrenia in a certain area of rural China, Ningxia Hui Autonomous Region.

In Chapter two, we used multi-level models to analysis the determinants of mental health service utilization in the patients with schizophrenia who had registered the community mental health systems. Determinants were categorized into multiple levels, personal, environmental, and health service. This study confirmed that the treatment gap of mental service had been significantly reduced in rural China, and concluded that individual characteristics were more related to the initiation of treatment while environmental and service characteristics were more relevant to maintain good treatment adherence.

Chapter three is meta-analysis to quantify the social deficits of patients with schizophrenia and identify the effect of both age and years of schooling on social cognition in healthy controls. By combining the RMET (Reading the mind in the Eyes Test) score of over 200 studies, we not only estimated the pooled the impairment of social cognition in patients with schizophrenia, but also modeled the non-linear change of social cognition over age using meta-regression with line construction (the social cognition peaks at around age of 30 and gradually decreases after it).

Chapter four collected empirical data of three groups of individuals, a group of drug-naïve patients with schizophrenia, a group of treatment controls, and a group of healthy controls. Beyond the direct comparison among three groups using fixed-effect models, we considered the completion status of scale as a proxy measurement

of general cognition and used Heckman estimation to adjust the large precent of uncompletion rate. This study confirmed that drug-naïve patients with schizophrenia had poorer social cognition then treated controls and appealed that we should design new tools to evaluate the social cognition for patients with lower level of education.

Chapter five concludes the whole thesis by summarising the main findings and future research directions.

Acknowledgments

I have to say that I almost knew nothing about economics before I came here. I spent six years obtaining my Bachelor's and Master's degrees in the School of Medicine, and another five years working in the field and mental hospitals before I registered as a Ph.D. student. Therefore, it could be imagined how hard the modules in the first year of Ph.D. were for me. I am very grateful to all the professors who taught these modules and especially my classmates. They were very patient and tried their best to help me to get into the study of economics. When I suddenly found myself able to follow the formulas and algorithm on the very heavy text book of Econometric and learn new models on my own after the first-year training, I was firstly so surprised, and then was appreciated the training provided by our school in both UK and China campus and my own persistence.

I am very grateful to all my four supervisors. They are Professor Chee Rury Hsieh, Professor Richard Hubbard, Professor Chen Zhuo, and Dr. Chew Chua Lian. They guiled me throughout my four-year Ph.D. study, and also provide me with enough freedom to choose my own research direction and questions.

Here, I also need to express my appreciation to Professor Michael R. Phillips. We have been working together for ten years since 2013 after I just graduated from my Master's degree. He supported me to interrupt my job to get a Ph.D. degree, provided me detailed advice for my field work to collect data, and also allowed me to use the data from his own project to complete this dissertation. For the article I published during my Ph.D., he did very detailed line by line editing for me.

Throughout the last three years of my Ph.D. study, people all over world were fighting with Covid-19 in their own. I am very grateful that all my family members

and friends have stayed healthy throughout these three years, which let me feel very relieved and being able to focus on my study.

Content

List of Tables

Chapter 1

Chapter 2

Chapter 3

Chapter 4

List of Figure

Chapter 2

Chapter 3

Chapter 4

Abbreviations

DALY	Disability-adjusted life year
DSM-IV	4th version of Diagnostic and Statistical Manual of Mental Disorders
DSM-5	5th version of Diagnostic and Statistical Manual of Mental Disorders
FIML	Full information maximum likelihood
GBD	Global Burden of Disease
HICs	High-income countries
LMICs	Low- and middle-income countries
MATRICS	Measurement and Treatment Research to Improve Cognition in Schizophrenia
MCCB	MATRICS Consensus Cognitive Battery
MCCB-CV-R	Revised Chinese version of MATRICS Consensus Cognitive Battery
mhGAP	Mental Health Gap Action Programme
MMSE	Mini-Mental State Examination
MSCEIT	Mayer-Salovey-Caruso Emotional Intelligence Test
NISPP	National Information System of Patients with Psychosis
PANSS	Positive and Negative Syndrome Scale
RMET	Reading the Mind in the Eyes Test
RMET-CV	Chinese version of Reading the Mind in the Eyes Test

RMET-CV-R	Revised Chinese version of Reading the Mind in the Eyes Test
SCID-IV	Structured Clinical Interview for DSM-IV
SMD	Standard mean difference
WHO	World Health Organization
WPA	World Psychiatric Association

Chapter One: Introduction

1.1 The disease burden of mental illness

In the twentieth century, a number of somatic health issues have been brought under control, for instance, communicable diseases, and the life expectancy have been improved dramatically. At the present time, non-communicable diseases and mental disorders become the next key issue in public health due to their enormous disease burden.

In the 1990s, epidemiologist and health economist from the World Bank, Havard School of Public Health, and the World Health Organization the first time developed a new metric, the disability-adjusted life year (DALY), to quantify the health effects of more than 100 diseases and injuries in multiple regions in the world, and commissioned the first Global Burden of Disease (GBD) Report (World Bank, 1993; Murray & Lopez, 1996; Murray, 1996). Under the monitoring of GBD study, the global number of DALYs due to mental illness increased from 80.8 million in 1990 to 125.3 million largely due to population growth and ageing, and the proportion of total DALYs attributed to mental illness also increased significantly, from 3.1% to 4.9% (GBD 2019 Mental Disorders Collaborators, 2022). Considering neurological disorder and substance use disorders, the disease burden of MNS (mental, neurological, and substance use disorders) explained 10.4% of total DALYs at ages combined (Patel et al., 2015). According to WHO, mental illnesses accounted for six of the 20 leading causes of disability in the worldwide in the population age between 15 to 44, the most productive age group.

The disease burden attribute to MNS is estimated to increase disproportionately in the coming decades, especially in developing countries due to a projected increase in the number of the young population entering the age of high risk for the onset of certain mental disorders (e.g., schizophrenia) and a further decrease of the disease burden attribute to infectious disease.

Besides the disease burden bore by the patients him or herself. The economic and social costs of mental disorders fall on a broader of people, not only people with mental

disorders and their caregivers, but also the government and even the whole society. Other than the direct cost of mental illness to pay for the targeted clinical treatment and intervention, e.g., drug cost and hospitalization, 'indirect cost' of lost productivity and premature death actually exceeded the direct cost found by most previous mental health economics studies (Hu, 2006; Chisholm et al., 2000; Patel et al, 2003; Murthy et al, 2005).

1.2 Mental health service and treatment gap

At the beginning of the 21st century, the WHO conducted the first global survey on mental health resources, the 'Atlas Project' (WHO, 2005; Saxena, Sharan, Garrido-Cumbrera, & Saraceno, 2006), and WHO updated the Atlas report in 2005, 2011, 2014, 2017 and 2020. Up to 2020, massive inequalities in the availability of mental health resources still existed, and so were the inequalities in resources allocated between high- and low-income countries and across regions. It also pointed out significant gaps all over the world between the existence of policies, plans, and laws and the service implementation (WHO, 2021).

Mental health services must be coordinated with other departments, such as social security, civil affairs, federation of disability, education by an adequate mental health policy or legislation. However, according to the WHO survey (shown in table 1.1, table 1.2, and table 1.3), 85% of the total responding countries had stand-alone mental health policies or plan and 65% of the countries had stand-alone mental health laws, but only 23% of the countries indicated the existence of indicators to monitor the implementation of mental health policy or law (WHO, 2020).

Table 1.1 Percentage of responding countries reporting the existence of stand-alone mental health policies or plans

	2014 (n=170)	2017 (n=175)	2020 (n=170)
WHO region			
AFR	71% (n=27)	72% (n=31)	76% (n=29)
AMR	80% (n=24)	82% (n=27)	91% (n=30)
EMR	67% (n=14)	78% (n=14)	80% (n=16)
EUR	79% (n=38)	81% (n=39)	91% (n=42)
SEAR	80% (n=8)	90% (n=9)	100% (n=8)
WPR	83% (n=20)	83% (n=19)	84% (n=21)

Global	77% (n=131)	79% (n=139)	85% (n=146)

Note: WHO Member States are grouped into six regions: Africa (AFR), the Americas (AMR), Eastern Mediterranean (EMR), Europe (EUR), South-East Asia (SEAR), and Western Pacific (WPR).

Source: WHO Mental Health Atlas 2020

Table 1.2 Percentage of responding countries reporting the existence of stand-alone mental health law

	2014 (n=158)	2017 (n=175)	2020 (n=171)
WHO region			
AFR	55% (n=18)	44% (n=19)	49% (n=19)
AMR	50% (n=14)	61% (n=20)	61% (n=20)
EMR	67% (n=12)	61% (n=11)	70% (n=14)
EUR	70% (n=33)	77% (n=37)	70% (n=32)
SEAR	60% (n=6)	50% (n=5)	63% (n=5)
WPR	73% (n=15)	83% (n=19)	80% (n=20)
Global	63% (n=99)	63% (n=111)	65% (n=111)

Note: WHO Member States are grouped into six regions: Africa (AFR), the Americas (AMR), Eastern Mediterranean (EMR), Europe (EUR), South-East Asia (SEAR), and Western Pacific (WPR).

Source: WHO Mental Health Atlas 2020

Table 1.3 Percentage of responding countries indicating existence of indicators/targets to monitor implementation of policies/plans (n=167)

	Indicators not available	Indicators available but not used	Indicators available and have been used in the past two years for a few components	Indicators available and have been used in the past two years for most or all components
WHO region				
AFR	22%	46%	27%	5%
AMR	34%	19%	22%	25%
EMR	15%	15%	35%	35%
EUR	20%	11%	40%	29%
SEAR	13%	25%	25%	38%
WPR	24%	16%	40%	20%
Global	23%	22%	32%	23%

Note: WHO Member States are grouped into six regions: Africa (AFR), the Americas (AMR), Eastern Mediterranean (EMR), Europe (EUR), South-East Asia (SEAR), and Western Pacific (WPR).

Source: WHO Mental Health Atlas 2020

The distribution of psychiatric service varies across different income countries. For instance, the median of psychiatric beds per 10,000 population in Europe was 8.00 while it was only 0.33 in South-East Asia (WHO, 2001). Furthermore, the distribution of psychiatric services setting also varies significantly between low- and middle-income countries and high-income countries. In South-East Asia, almost all the psychiatric beds (83%) are located in mental hospitals compared with only 64% in Europe. Though community-based mental service is well recognized as the most effective way to promote mental health, having a better effect than institutional treatment settings on the outcome and quality of life on patients with chronic mental illness, the development of community mental health service in low-income countries are still far behind. Overall, 0.64 community-based mental health facilities exist per 100,000 population in all the responding countries to the WHO atlas survey, but there is extreme variation between income groups of the countries, with 0.11 facilities per 100,000 in low-income countries and 5.1 facilities per 100,000 population in high-income countries (more than 50 folder difference).

As a consequence of the inequality of mental health service provision, together with some other factors (e.g., stigma, mental health literacy), a large proportion of the individuals suffering from any mental illness have not yet received any mental health care, i.e., treatment gap. Based on the WHO report, the treatment gap of severe mental illness (e.g., schizophrenia) and mood disorders (depressive disorder and bipolar disorder) was about one-third and 50% respectively, and treatment gap of substance use disorder (e.g., alcohol use disorder) exceeded three-fourth (WHO, 2001).

Reference:

Chisholm, D., Sekar, K., Kumar, K.K., Saeed, K., James, S., Mubbashar, M., & Murthy R.S. (2000). Integration of mental health care into primary care. Demonstration cost-outcome study in India and Pakistan. *British Journal of Psychiatry*, 176, 581-588. Doi: 10.1192/bjp.176.6.581

GBD 2019 Mental Disorders Collaborators. (2022). Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet*, 9(2), 137-150. Doi: http://doi.org/10.1016/S2215-0366(21)00395-3

Hu, T.W. (2006). Perspectives: an international review of the national cost estimates of mental illness, 1990-2003. Journal of Mental Health Policy and Economics, 9(1), 3-13.

Murthy, R.S., Kumar, K.V.K., Chisholm, D., Thomas, T., Sekar, K., & Chandrashekari, C.R. (2005). Community outreach for untreated schizophrenia in rural India: a follow-up study of symptoms, disability, family burden and costs. Psychological Medicine, 35(3), 341-351. Doi: 10.1017/s0033291704003551.

Patel, V., Chisholm, D., Rabe-Hesketh, S., Dias-Saxena, F., Andrew, G., & Mann, A. (2003). Efficacy and cost-effectiveness of drug and psychological treatments for common mental disorders in general health care in Goa, India: a randomized, controlled trial. The Lancet, 361(9351), 33-39. Doi: 10.1016/S0140-6736(03)12119-8.

Patel, V., Chisholm, D., Parikh, R., Charlson, F.J., Degenhardt, L., Dua, T., … Whiteford, H. (2015). Addressing the burden of mental, neurological, and substance use disorders: key messages from Disease Control Priorities, 3rd edition. *The Lancet*, 387(10028), 1672-1685. Doi: https://doi.org/10.1016/S0140-6736(15)00930-6

Murray CJL, Lopez AD, editors (1996). The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020. Cambridge: Harvard University Press.

Murray CJL (1996). Rethinking DALYs. In: Murray CJL, Lopez AD, editors (1996). The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020. Cambridge: Harvard University Press.

Saxena, S., Sharan, P., Garrido-Cumbrera, M., & Saraceno, B. (2006). World Health Organization's Mental Health Atlas 2005: implications for policy development. World Psychiatry, 5(3), 179-184.

World Bank (1993). World Development Report 1993. Washington: World Bank.

World Health Organization. (2001). Mental Health: New Understanding, New Hope. World Health Organization, Geneva.

World Health Organization. (2005). Mental health atlas 2005. World Health Organization, Geneva.

World Health Organization. (2021). Mental health atlas 2020. World Health Organization, Geneva.

Chapter Two: Relationship of the characteristics of patients and the mental health service network to the care-seeking behavior and antipsychotic medication adherence of community-dwelling individuals with schizophrenia in northwest China: a multilevel-model analysis

Abstract

Introduction: schizophrenia is quite a prevalent severe mental illness, but the treatment gap among patients with schizophrenia widely exists in low- and middle-income countries. China started its natiowide community mental health service program in 2004. However, the treatment gap under this program remained unclear.

Methods: A cross-sectional survey covering all the registered patients with psychosis in the community mental health service system in eight counties in Ningxia Hui Autonomous Regions. Patients were interviewed about their demographic information, disease history, and service utilization. The characteristics of each village or urban community were also collected to link with the patients in patients living in each geographic unit.

Results: In total, 2280 patients with schizophrenia were enrolled in this study. Among these patients, 93.5% ever started antipsychotic treatment, 74.0% ever used antipsychotic medicines in the three months prior to the survey, and 64.1% had good adherence to drug use in these three months. Not only personal factors were related to medication use, environment characteristics (closer to mental hospitals, living in a area with less ethnic minority population) were also related to better mental health service utilization.

Conclusion: The community mental health program have largely eliminate the treatment gap among patients with schizophrenia in rural China in the recent years. However, the inequity of mental service provision and utilization across different areas

still exists significantly.

2.1 Introduction

Schizophrenia is a severe mental illness with a lifetime prevalence of about 1% in China (Phillips et al., 2009; Moreno-Kustner, Martin, & Pastpr, 2018) that typically starts during young adulthood (peak incidence between 15 to 35 years of age) and usually results in chronic disability. As estimated by Global Burden of Disease study, schizophrenia accounts for about 1% of the total global economic and social burden attributed to all diseases and injuries and about 1.7% of total YLDs (years of life lived with disability) (Charlson et al., 2018), resulting in an annual loss of 12.7 million (95% uncertainty interval 9.5 to 15.6 million) disability-adjusted life years (DALYS) (He et al., 2020).

A report from the World Health Organization (WHO, 2013) suggests that the excess disability associated with schizophrenia is partially the result of delays in the initiation of appropriate treatment. This delay can last for years (Blanc, Newman, & Orn, 1997); some individuals suffering from this severe, chronic illness *never* receive appropriate treatment (Kohn, Saxena, Levav, & Saraceno, 2004). Moreover, many individuals with schizophrenia who start treatment do not stay on treatment. The 'treatment gap' for psychotic disorders – the percentage of individuals with psychotic conditions such as schizophrenia who are not receiving appropriate treatment – is an important metric of the coverage and quality of a community's mental health care system in both low- and middle-income countries (LMICs) and high-income countries (HICs) (WHO, 2008; Thornicroft, Szmukler, Mueser, & Krake, 2012).

The 12-month treatment gap for schizophrenia (that is, the % of all individuals with schizophrenia who do not receive appropriate antipsychotic medication treatment for the entire prior 12-month period) reported in 7 HICs ranged from 5.9% to 61.5% (median, 26.9%)(Kohn et al., 2004). More recent reports from HICs suggest that the delay in treatment initiation has decreased over time and patients' adherence to prescribed medication has increased over time (possibly due to promulgation of long-acting injectable antipsychotic medications, behavioral interventions, and technology-

supported interventions) (Curto et al., 2020). For example, in Singapore only 19.4% of individuals who developed a psychotic disorder in 2016 had not consulted a mental health professional within one year of the onset of the illness (Subramaniam et al., 2021), and papers from HICs published in 2010 or later report adherence rates with oral or intramuscular antipsychotic treatment among patients with schizophrenia between 68% and 81% (Lin et al., 2021; Kishi, Oya, & Iwata, 2016; Gutierrez-Casares et al., 2010).

The situation is less hopeful in LMICs. A survey about access to and utilization of mental health services of patients with schizophrenia in 50 LMICs from 2005 to 2010 using the WHO's 'Assessment Instrument for Mental Health Systems' (Lora et al., 2012) reported median treatment gaps in the 12 months prior to the survey of 89% in low-income countries, 69% in lower-middle-income countries, and 63% in upper-middle-income countries. (The treatment gap in China was 59%.) As expected, the survey found that the availability of psychiatric services was a significant predictor of the treatment gap in the 50 countries. In an attempt to address this issue, the WHO's Mental Health Gap Action Programme (mhGAP) in LMICs (WHO, 2008) includes recommendations to promote interventions focused on overcome existing obstacles to care-seeking. However, undertaking such interventions requires country-specific, region-specific, and community-specific information about the current treatment gap for psychotic disorders and about the relative importance of different barriers to seeking care in targeted populations – information that is rarely available.

From an economic standpoint, reasons for the treatment gap fall into two categories: demand-side factors and supply-side factors. Demand-side factors include the stigma and discrimination associated with mental illness, lack of mental health literacy, and poor adherence to prescribed medications; supply-side factors include the limited number of mental health professionals, availability of antipsychotic medications, and disparities in the coverage and quality of mental health services.

The 'treatment rate' (the inverse of the treatment gap) for schizophrenia is a macro-level indicator for a specific population cohort or geographic region that is defined as the proportion of all patients in the cohort that take antipsychotic medication as

prescribed. This rate is associated with patients' access to appropriate mental health care services, but – somewhat surprisingly, previous studies have found that patient-related factors are *not* consistently associated with adherence to prescribed antipsychotic medication (Fenton, Blyler, & Heinssen, 1997; Thornicroft et al., 2011). For example, based on the WHO accessibility survey in 12 LMICs countries and 12 HICs (Thornicroft et al., 2011), in most countries women with schizophrenia were more likely to adhere to treatment than men (China was the only exception), and a significant positive association between a patient's income level and their use of mental health services was only identified in a few of the participating counties (including Colombia, China, Israel, and Japan). One possible explanation for the inconsistency in results about the relationship of personal characteristics to the utilization of mental health services is that the analyses are not adequately controlled for differences in the availability of services. Moreover, the heterogeneity of the availability of services within countries (particularly between urban and rural locations) is obscured when comparing service utilization between countries, because such comparisons necessarily use national-level macro measures.

One way to address this issue is to consider the relationship of personal-level factors to the utilization of mental health services in countries where there are measurable differences in access to mental health services between different regions. Based on a 2018 report in the Lancet, in 2016 China had the largest provincial-level gap in the coverage and quality of health services of any country in the world (GBD 2016 Healthcare Access and Quality Collaborators, 2018), so China is an ideal location to identify possible differences in the relevance of person-level factors in locations with different levels of access to services. As of the end of 2015, 42.4% (1180/2783) of all counties in the country had no mental health services whatsoever (Shi et al., 2019). Most mental health professionals in China live in urban areas in the more affluent eastern provinces: there is a 5-fold difference in the number of psychiatrists per 100,000 population between China's provincial-level administrative districts (from 0.88/100,000 in Qinghai Province to 5.57/100,000 in Beijing) and a 20-fold difference in

the number of psychiatric beds per 10,000 population (from 0.37/10,000 in Qinghai Province to 6.2/10,000 in Beijing)[1] (Shi et al., 2019). It is unknown whether these huge differences in the availability of mental health providers and services across the country affect the relationship of individual-level factors (such as gender, age, education, and income) with care-seeking for mental health services. To our knowledge, there has been no research that specifically aims to identify the interrelationships linking environmental factors (i.e., availability of services), patient-related factors, and care-seeking for mental health services in China.

Two factors stimulated the development of community-based mental health services in developed countries in the second half of the 20th century: the discovery of effective antipsychotic medications, and community concern about the negative effects of long-term (often lifetime) confinement in chronic psychiatric hospitals (Thornicroft et al., 2011). This inclusion of community-based services as part of the mental health delivery system has gradually been adopted in most LMICs. The WHO Mental Health Action Plan for 2013-2020 (WHO, 2013) considered community-based mental health services an essential component of the effort to reduce the treatment gap for psychosis. Both the World Psychiatric Association's (WPA) and the World Bank also highlight the central role community-based mental health care should play in a comprehensive package of mental health services: they recommend that community services should be responsible for early case finding and identification, initiation of treatment, specialist supervision of community-based services, promotion of self-help or peer support, community-based rehabilitation, and collaboration with other social welfare organizations. (Thornicroft et al., 2011; Thornicroft, Deb, & Henderson, 2016)

In parallel with this global trend to expand the provision of community-based mental health services, in 2004 the Chinese Health Ministry initiated the National Information System of Patients with Psychosis (NISPP) as the main platform for providing community-based mental health services to persons with severe mental disorders (Ma et al., 2009). Services provided by the network include the identification and diagnosis

[1] As of 2015, Tibet had no psychiatrists and no psychiatric beds in the entire province.

of patients; provision of free medication and, in some cases, free hospitalization; three-monthly follow-up of patients by local non-psychiatric healthcare workers; and free regular physical examinations that include blood tests which screen for side effects of antipsychotic medication. By the end of 2019, all 333 cities and 2854 counties in the country were covered by the NISPP. At that time there were 6,230,157 patients actively managed by the registry, 72.1% of whom had a diagnosis of schizophrenia (Zhao et al., 2020). This is the largest community mental health service network in the world. However, given the unequal distribution of mental health professionals and mental health services around the country, the substantial financial resources provided to the NISPP by the national and provincial governments (by the end of 2015, the cumulative investment exceeded 2.2 billion RMB – about 0.35 billion USD) are disproportionately targeted to locations that already have a mental health infrastructure that can use the funds to expand their current services while excluding locations without any mental health infrastructure. Thus, this very substantial national-level financial and administrative initiative may be increasing rather than decreasing the inequitable distribution of mental health services in the country. To date, there has been no systematic evaluation of the effectiveness of the NISPP program that could help resolve this important issue and, possibly, provide strategies for improving the equitability of the program.

This report aims to help bridge the gap in our knowledge about this core component of China's mental health care delivery system. We present the results of a comprehensive survey of the NISPP program in Ningxia Hui Autonomous Region (a province-level administrative region) based on telephone interviews with the guardians of a representative sample of 2280 patients with schizophrenia registered in the system from 2016 to 2018. The data collected in this survey were used in a multilevel modeling procedure that 1) identified the interactive relationships between patient-level determinants, the local availability of mental health personnel and services, and care-seeking by patients and their family members; and 2) assessed the effectiveness and equity of the NISPP program in Ningxia.

2.2 Methods

2.2.1 Participants and settings

We conducted a cross-sectional survey that interviewed the guardians of registered patients in the NISPP system in Ningxia Hui Autonomous Region, China (shown in Figure 2.1). The survey was supported by Shanghai Jiaotong University, Columbia University, Harvard University, New York University, and Ningxia Medical University. Ningxia is a province-level area in northwest China in which about one-third of residents are members of the Hui ethnic minority and two-thirds are members if the Han ethnic group (the main ethnic group in China). The average GDP per capita in Ningxia in 2019 was 7858 USD (which is close to the median GDP per capita across China's 31 provinces of 8189 USD).The number of health professionals per 1000 population in Ningxia is greater than the average level in mainland China (6.6 vs 6.1 in 2016) (National Health Commission of the People's Republic of China, 2017); however, as of 2015 both the mental health manpower (1.51 psychiatrists per 100,000 population) and the availability of psychiatric beds (1.52 per 10,000 population) in Ningxia was ranked 28th among mainland China's 31 provinces (Shi et al., 2019). As is true throughout China, most psychiatrists in Ningxia work in specialized psychiatric hospitals; there were only two general hospitals in the province that provided psychiatric services (both in Yinchuan, the provincial capital) at the time of the survey (2016-2018), and both of them only provided outpatient services.

Figure 2.1 Location of Ningxia (total area 66,000 squared kilometer, population 6,301,350, 55% urban population, 34% Muslim)

We selected 8 representative counties from among the 22 counties in Ningxia: two primarily urban counties with a high proportion of Han ethnicity residents, two primarily urban counties with a high proportion of Hui ethnicity residents, two primarily rural counties with a high proportion of Han ethnicity residents and two primarily rural counties with a high proportion of Hui ethnicity residents. All individuals registered in the NISPP in these 8 counties with a recorded diagnosis of schizophrenia were potential respondents in the survey. Inclusion criteria (assessed by a trained interviewer who interviewed the patient's guardian by phone) were:

a) A history of 'positive' psychotic symptoms including hallucinations, delusions, thought disorder, or grossly bizarre behavior. [In 24.2% of the registered patients with a diagnosis of schizophrenia, the diagnosis was made by a local, non-psychiatric health care worker (usually a public health doctor or a village

doctor); so to reduce the likelihood of including individuals who do not meet the formal diagnostic criteria of schizophrenia in the analysis, we exclude those for whom there is no reported history of these cardinal symptoms of schizophrenia.]

b) No history of mental retardation, epilepsy, brain injury or serious brain disease.

c) Patient's guardian provides verbal agreement to participate in the survey.

As shown in Figure 2.2, among the 7,141 active registrants in the NISPP in the eight selected counties, 2,280 met the inclusion criteria. These individuals were distributed throughout the 141 rural townships and urban sub-districts (further subdivided into 990 rural villages and urban neighborhoods) in these 8 counties.

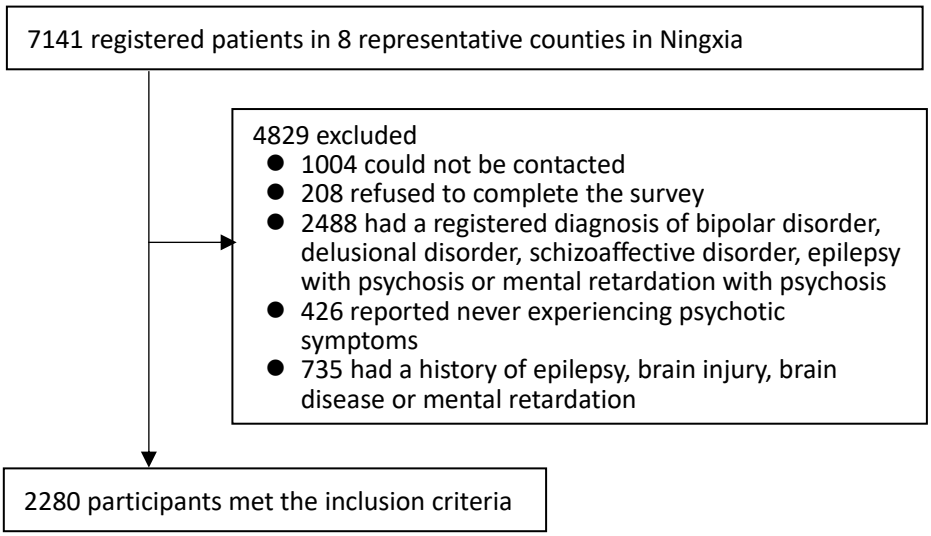

Figure 2.2 Flowchart of the participants' enrollment

2.2.2 Data collection

The individual-level data was collected by telephone interviews with patients' guardians (n=2108) or – in a small number of cases in which a telephone survey was not possible (n=172) – by face-to-face interviews in the patients' homes or the local health

clinic. Trained research staff from Shanghai Jiaotong University conducted most of the telephone surveys from June 2016 through May 2018; about one-third of the telephone surveys were conducted by fourth and fifth year undergraduate public health students and medical students who were trained and supervised closely by the research staff. Community-based health professionals helped the research staff identify and interview guardians who could not be contacted by phone.

We also interviewed about 1000 rural village doctors and urban neighborhood doctors in the 8 counties. They provided village-level (rural communities) or neighborhood-level (urban communities) information about community characteristics: number of Han-ethnicity versus minority-ethnicity residents, and the number residents in each community living in households that are officially classified as a 'rural household' or an 'urban household'.

2.2.3 Dataset

Dependent variables of interest:

1) ***life-time treatment-seeking behaviour***: a binary variable indicating whether or not the patient EVER received (and used) treatment with antipsychotic medication;

2) ***any use of antipsychotic medication in the previous 3 months***: a binary variable indicating whether or not the patient had used antipsychotic medication at any time over the three months prior to the interview;

3) ***adherence to treatment with antipsychotic medication in the previous 3 months***: a binary variable indicating whether or not the patient regularly used antipsychotic medication over the three months prior to the interview (i.e., on at least 80 days).

Independent variables of interest: this study collected both individual-level and cluster(community)-level (i.e., rural village or urban neighborbood) measures

(hereafter 'community' or 'cluster' refers to the combined category of rural villages and urban neighborhoods):

1) ***individual-level variables***: age, gender, years of formal schooling, place of residence (urban v. rural), ethnicity (Han v. ethnic minority), marital status (currently married v. not currently married), whether patient lives alone (i.e., without family members), guardian's relationship with patient, year of onset of illness, year and month of starting antipsychotic medications, whether the patient had ever been an inpatient in a psychiatric hospital (or ward), whether the patient received free antipsychotic medication from the NISPP system at the time of the survey, whether the patient received free medication from the Disabled Person Federation (DPF) at the time of the survey, whether the patient received any follow-up monitoring by a primary care provider in the prior year, whether patient is registered as disabled in the DPF network (most, but not all, persons registered as disabled receive monthly benefits), whether receiving monthly subsidy due to poverty from civil affairs department, and whether patient's family is registered as an impoverished household at time the patient was enrolled in the NISPP service system[2] ;

2) ***community (cluster)-level variables***: per capita GDP of the county in 2016 (extracted from 2017 Ningxia Statistical Yearbook), whether there were mental health professionals working in the county of residence when the survey was conducted, travel distance from patient's home to the nearest psychiatric hospital at the time of the interview (derived from Gaode Map app in 2020 July), per capita medical professionals (of any type) in the county, whether the county is mountainous[3], number of patients with diagnosis of schizophrenia registered in the NISPP per 1000 population in the county, percentage of residents living in the community who are registered as a member of a rural household (the

[2] In 2016, the national criteria for impoverished household was mean per capita income < 2952 RMB (about 450 USD).

[3] Ningxia province is divided into mountainous and non-mountainous regions. Three of the 8 counties surveyed (Xiji, Pengyang, and Haiyuan) were located in mountainous regions.

individual's 'hukou' registration; hereafter these individuals are referred to as 'rural residents'), percentage of ethnic minority residents living in the community, and characteristics of rural village-level or urban neighborhood-level primary care provider responsible for monitoring the care of the patient (gender, age, educational attainment, work experience, etc.),

2.2.4 Modelling strategy

Patients in each community are monitored by the same primary care provider (either a village doctor or an urban neighborhood doctor), so in the analyses rural villages and urban neighborhoods were considered 'clusters' (also labelled 'communities') and multilevel modelling was used to adjust for any potential correlation of results among patients in the same cluster/community. Two-level random-intercept regression (level 1: individual, level 2: community) is used to link the characteristics of the rural villages and urban neighbourhoods to the characteristics of individual patients.

All the analysis is conducted using STATA version 16.

2.2.4.1 <u>Outcome One</u>: Ever use antipsychotic medication in lifetime

A binary outcome variable y_i (y_i=1 if the patient i ever used antipsychotic medication, y_i=0 if patient i never used antipsychotic medication) is defined and a two-level random-intercept logistic regression is used to model the probability that the patient ever used antipsychotic medication:

$$logit\left[Pr\left(y_{ij}=1\middle|X_{ij},\zeta_j\right)\right]=\left(\zeta_j+W'\gamma_e\right)+\beta_{e0}+\left(\delta_{e1}W_{r\,j}+\delta_{e2}W_{m\,j}+\delta_{e3}W_{d\,j}\right)\ years\ of\ schooling_{ij}+\left(\theta_{e1}W_{r\,j}+\theta_{e2}W_{m\,j}+\theta_{e3}W_{d\,j}\right)age_{ij}+X'\beta_e+\varepsilon_{ij}\quad(1)$$

where X_{ij} is a vector of covariates including characteristics of both patient i living in community j, and patient-level characteristics include age, years of schooling, gender, marital status, ethnicity, year of onset of psychotic symptoms, and whether patient's

family is registered as an impoverished household. Community-level characteristics, represented by the vector W_j, include distance from the community to the nearest psychiatric hospital, whether the area is mountainous, percentage of rural residents in the community, percentage of ethnic minority residents in the community, whether psychiatric services are provide within the county, and the average GDP per capita in the county. Given that there may be a non-linear relationship between distance to the nearest psychiatric service and the outcome, the squared term of distance to hospital was also included in the W_j vector.

Ideally we would like to fit a random-coefficient model to assess the possibility that patients' current age and years' of schooling could have differential effects on the treatment received for patients living in rural versus urban communities and at different distances from a psychiatric facility. However, the sample size was not sufficient to fit a random-coefficient model. As a more feasible alternative, we used a random-intercept model that includes six interaction terms which makes it possible to generate community-specific coefficients for age and years of schooling: 1) patient's years of schooling and percent rural residents in the community; 2) patient's years of schooling and percent ethnic minority residents in the community; 3) patient's years of schooling and distance to nearest psychiatric hospital; 4) patient's age and percent rural residents in the community; 5) patient's age and percent ethnic minority residents in the community; and 6) patient's age and distance to nearest psychiatric hospital.

In the random-intercept model shown in equation (1) W_{rj} is the percent of rural residents in community j, W_{mj} is the percent of ethnic minority residents in community j, and W_{dj} is the distance to the nearest psychiatric hospital from community j (these three community-level variables are also included in the composite community-level variable W_j); the corresponding coefficients for the six interaction terms are δ_{e1}, δ_{e2}, δ_{e3}, θ_{e1}, θ_{e2}, and θ_{e3}. ζ_j ~N (0, ψ) is the random intercept of community j; this parameter is assumed to be identically distributed across the patients in community j and independent of the covariates vector X (individual-level information) and W (community-level information). The sum of $W'\gamma_e$ and ζ_j can be considered the

community(cluster)-specific effect on the outcome (i.e., ever receiving antipsychotic medication) after controlling for the individual-level covariates in X.

2.2.4.2 <u>Outcome Two</u>: Any use of antipsychotic medication in the previous three months among patients who had ever used antipsychotic medication

A binary outcome variable, Z_{ij}, denotes the use of antipsychotic medication in the three months before the survey of patient i in community j (Z_{ij}=1 denoting ever using antipsychotic medication in the prior three months, and Z_{ij}=0 denoting not using any antipsychotic medication in the prior three months). In this outcome the regularity of medication use over the last three months is *not* considered; what is important is whether the patient used antipsychotic medication at ANY time during the prior three months. Among patients who had EVER used antipsychotic medications at any time over the entire course of their illness (excluding patients who had never used antipsychotic medications), the probability of using medication at any time in the three months prior to the interview is modelled by equation 2.

$$logit\left[Pr\left(Z_{ij}=1\middle|X_{ij},\eta_j\right)\right]=\left(\eta_j+W'\gamma_c\right)+\beta_{c0}+\left(\delta_{c1}W_{r\,j}+\delta_{c2}W_{m\,j}+\delta_{c3}W_{d\,j}\right)*years\ of\ schooling_{ij}+\left(\theta_{c1}W_{r\,j}+\theta_{c2}W_{m\,j}+\theta_{c3}W_{d\,j}\right)*age_{ij}+X'\beta_c+\varepsilon_{ij} \quad (2)$$

η_j ~N (0, μ) is the random intercept of community j; this parameter is assumed to be identically distributed across the patients in community j and independent of the covariate vectors X and W. The patient-level characteristics variable (X) includes all the characteristics considered in equation (1) and one additional covariate: whether the patient had ever been an inpatient in a psychiatric hospital (or ward). The community-level characteristics variable (W) includes the same community covariates considered in equation (1). The six interaction terms included in equation (1) are also included in this equation: years of schooling*percent of rural residents, years of schooling*percent of ethnic minority residents, years of schooling*distance to the nearest psychiatric

hospital, age*percent of rural residents, age*percent of ethnic minority residents, and age*distance to the nearest psychiatric service; the corresponding coefficients for these interaction terms are δ_{c1}, δ_{c2}, δ_{c3}, θ_{c1}, θ_{c2}, and θ_{c3}.

2.2.4.3 <u>Outcome Three</u>: Regular use of antipsychotic medication in previous three months among patients who had ever used antipsychotic medication in the previous three months

A binary outcome variable, R_{ij}, denotes the current medication usage of the patient i in community j (R_{ij}=1 denoting 'regularly' using antipsychotic medication in the prior three months, and R_{ij}=0 denoting not regularly using medication in the prior three months). 'Regularly using medication' is defined as using oral antipsychotic medication for at least 80 days in the three months prior to the interview or using long-acting injectable antipsychotic medications as prescribed by a physician in the 3 months prior to the interview. Random-intercept multilevel logistic regression was used to estimate the probability of regularly using medication over the three months prior to the survey ($R_{ij} = 1$) among patients who had taken antipsychotic medication at ANY time in the prior three months.

$$logit\left[Pr\left(R_{ij} = 1 \middle| X_{ij}, \kappa_j\right)\right] = \left(\kappa_j + W'\gamma_r\right) + \beta_{c0} + \left(\delta_r W_{r\,j}\right) age_{ij} + X'\beta_r + \varepsilon_{ij} \qquad (3)$$

κ_j ~N (0, ϕ) is the random intercepts of community j; this parameter is assumed to be identically distributed across the patients in community j and independent of the covariate vectors X and W. The variable incorporating patient characteristics (X) includes all the covariates consider in equation (2) and two additional covariates: whether the patient was receiving free antipsychotic medication from the NISPP system at the time of the survey, and whether the patient was receiving free medication from the Disabled Person Federation (DPF) at the time of the survey. Only one interaction term, percentage of rural residents*age is included in this equation.

2.2.5 Prediction

To demonstrate the influence of community(cluster)-level factors, we estimate the community-specific probability of ever receiving treatment and the community-specific probability of using antipsychotic medication at any time in the prior 3 months. To achieve this, we first estimate the community-specific value of the random-intercept in each community (ζ_j in equation 1 and η_j in equation 2) using the Empirical Bayes Prediction method (Rabe-Hesketh, 2012). The community-specific probabilities for the two outcomes of interest are then estimated by entering the estimated community-specific intercept, specifying the individual-level predictors as their mean values in the entire sample (shown in Table 2.1) and entering the community-specific values for the community-level variables in equations 1 and 2, respectively. After estimating the predicted probability of each of the two outcomes in each community, we then calculate the predicted probabilities for each rural township and urban sub-district (administrative units including multiple rural villages or urban neighborhoods, respectively) by taking the mean of probability of all the communities included the corresponding rural township or urban sub-district; these probabilities are then displayed on a provincial map of Ningxia using Arcgis (version 10.2).

2.3 Results

2.3.1 Characteristic of the participants and the communities in which they reside

The basic characteristics of the 2280 participating patients and the 990 communities they lived in at the time of the survey are shown in Table 2.1. The educational level of participating patients was quite low: 441 (19.4%) had never attended school and 1018 (44.9%) had less than 6 years of formal schooling. These community-dwelling patients with schizophrenia had a duration of illness ranging from 1 to 74 years (mean 20.6 years). Only 148 (6.5%) had never used antipsychotic medication, 1684 (74.0%) had used antipsychotic medication at some point in the prior three months, and 1446 (64.1%) had used antipsychotic medication regularly over the prior three months. As is

the case for many parts of northwestern China, the 8 representative countries from Ningxia included in the survey had very limited mental health services: 89.6% of the communities (887/990) were parts of counties that did not have a psychiatric hospital or any specialized psychiatric services in a general hospitals.

Table 2.1 Characteristics of 2280 patients with schizophrenia enrolled China's National Information System of Patients with Psychosis in Ningxia Province and of the 990 communities in which they live

Individual-level characteristics of patients with schizophrenia (n=2280)	
Female (n, %)	1,290 (56.6%)
Age in years (mean [sd])	49.4 (12.8)
Member of ethnic minority group (n, %)	687 (30.1%)
Years of formal schooling (mean [sd]) [a]	6.6 (4.7)
Currently married (n, %) [b]	1,297 (56.9%)
Patient's household is classified 'rural household' by the government (n, %)	1553 (68.1%)
Patient's household is classified 'impoverished' by the government (n, %) [c]	1,863 (81.7%)
Duration of illness since first onset of psychotic symptoms in years (mean [sd]) [d]	20.6 (11.1)
Duration of untreated psychosis prior to initiation of antipsychotic treatment in months (median [IQR]) [e]	1 (1-12)
Ever received treatment with antipsychotic medications (n, %)	2,132 (93.5%)
Used antipsychotic medication at any time in the last three months (n, %)[f]	1,684 (74.0%)
Regularly used antipsychotic medication in last three months (n, %)[g]	1,446 (64.1%)
Currently receiving free antipsychotic medication (n, %)[h]	1,183 (52.5%)
Currently receiving free antipsychotic medication from NISPP system (n, %)[i]	862 (38.2%)
Currently receiving free antipsychotic medication from DPF system (n, %)[j]	404 (17.9%)
Community-level characteristics (n=990)	
Percentage of residents living in households officially registered as 'rural household' (mean [sd]) [k]	89.2% (27.3%)
Percentage of non-Han ethnic minority residents (mean [sd]) [l]	47.0% (43.3%)
Distance of administrative center of community to the nearest psychiatric hospital in km (mean, sd)[m]	74.6 (38.7)
Community is part of a county that has a psychiatric hospital (n, %)	103 (10.4%)
Community is part of a county officially classified as 'mountainous' (n, %)	450 (46.2%)
GDP per capita of the county at time of survey in RMB (mean [sd])	36122 (32586)

[a] 11 cases missing years of schooling
[b] 18 cases missing marital status
[c] 166 cases missing classification of household status
[d] 40 cases missing duration of illness
[e] 92 cases missing duration of untreated psychosis; if never treated, this duration is the same as the duration of illness
[f] 4 cases missing information about whether used antipsychotic medication at any time in the last three months
[g] 23 cases missing information about whether regular used antipsychotic medication in the last three months
[h] 25 cases missing information about whether receiving free medication in the last three months
[i] 23 cases missing information about whether receiving free medication from NISPP
[j] 25 cases missing information about whether receiving free medication from DPF (Disabled Person Federation)
[k] 12 communities missing percent of residents living in rural households
[l] 16 communities missing percent of non-Han ethnic minority residents
[m] 2 communities missing distance to the nearest psychiatric hospital

2.3.2 Ever used antipsychotic medication

Overall, 93.5% of the patients had received antipsychotic medication at some point in their lifetime. The results of the multilevel regression analysis (equation 1) that assessed the probability of any lifetime use of antipsychotic medication is shown in Table 2.2.

Table 2.2 Multi-level analysis of ever using antipsychotic medication (n=2,028)[a]

	coefficient	OR (90% CI)	p
Individual-level characteristics			
gender (1=female 2=male)	**-0.716**	**0.489 (0.331-0.722)**	**0.003****
age in years	-0.011	0.989 (0.939-1.041)	0.717
ethnicity (1=Han 2=ethnic minority)	-0.510	0.600 (0.339-1.062)	0.141
Patient's household is classified 'rural household' (1=rural, 2=urban)	0.143	1.154 (0.470-2.833)	0.794
years of formal schooling	0.078	1.081 (0.933-1.253)	0.385
marital status (1=currently married, 2=other)	-0.081	0.923 (0.602-1.414)	0.756
patient's household classified as 'impoverished' (1=yes, 2=no)	-0.252	0.778 (0.386-1.568)	0.555
calendar year of initial onset of psychotic symptoms	**0.022**	**1.023 (1.004-1.041)**	**0.040***
Community-level characteristics			
percent rural residents	**-3.875**	**0.021 (0.001-0.772)**	**0.078***
percent ethnic minority residents	1.747	5.736 (0.606-54.258)	0.201
distance to the nearest psychiatric hospital (km)	-0.004	0.996 (0.955-1.038)	0.860
quadratic term of distance to the nearest psychiatric hospital (km)	**0.0002**	**1.000 (1.000-1.0003)**	**0.096***
average GDP per capita of the county at time of survey	0.084	1.088 (0.972-1.217)	0.217
community is part of a county officially classified as 'mountainous'	0.096	1.101 (0.682-1.777)	0.742
has psychiatric hospital within the county (1=yes 2=no)	0.514	1.673 (0.787-3.553)	0.261
Interaction terms			
years of schooling*percent rural residents in community	0.088	1.092 (0.939-1.271)	0.338
years of schooling*percent ethnic minority residents in community	-0.090	0.914 (0.810-1.030)	0.216
years of schooling*distance to the nearest psychiatric service	<0.001	1.000 (0.999-1.001)	0.997
age*percent rural residents in community	0.052	1.053 (0.997-1.112)	0.119
age*percent ethnic minority residents in community	**-0.043**	**0.958 (0.924-0.993)**	**0.051***
age*distance to the nearest psychiatric service	**-0.001**	**0.999 (0.999-1.000)**	**0.057***

[a] 252 cases were removed from the analysis because they have missing values in one or more covariates
Rho (residual intraclass correlation) = 0.253; Sigma_u =1.055 (95% CI: 0.653-1.705)
* p<0.1, ** p<0.01, *** p<0.001

As shown in Table 2.2, female gender, later calendar year of the onset of illness, and a high proportion of urban residents in the village/neighborhood and the quadratic term

of distance to the nearest psychiatric hospital were significantly (at the p<0.10 level) positively related with a history of ever having used antipsychotic medication.

As shown in Figure 2.3, after adjusting for other variables in the model, the relationship of distance to nearest psychiatric hospital and probability of lifetime use of antipsychotic medication is different for different ranges in the distance to the nearest psychiatric hospital. At distances of under 90 km, greater distance to psychiatric hospital decreases the likelihood of any lifetime use of antipsychotic medication; however, at distances of over 90 km, greater distance to psychiatric hospital *increases* the likelihood of lifetime use of antipsychotic medication, an effect that becomes statistically significant (at the P<0.10 level) for distances of greater than 150 km.

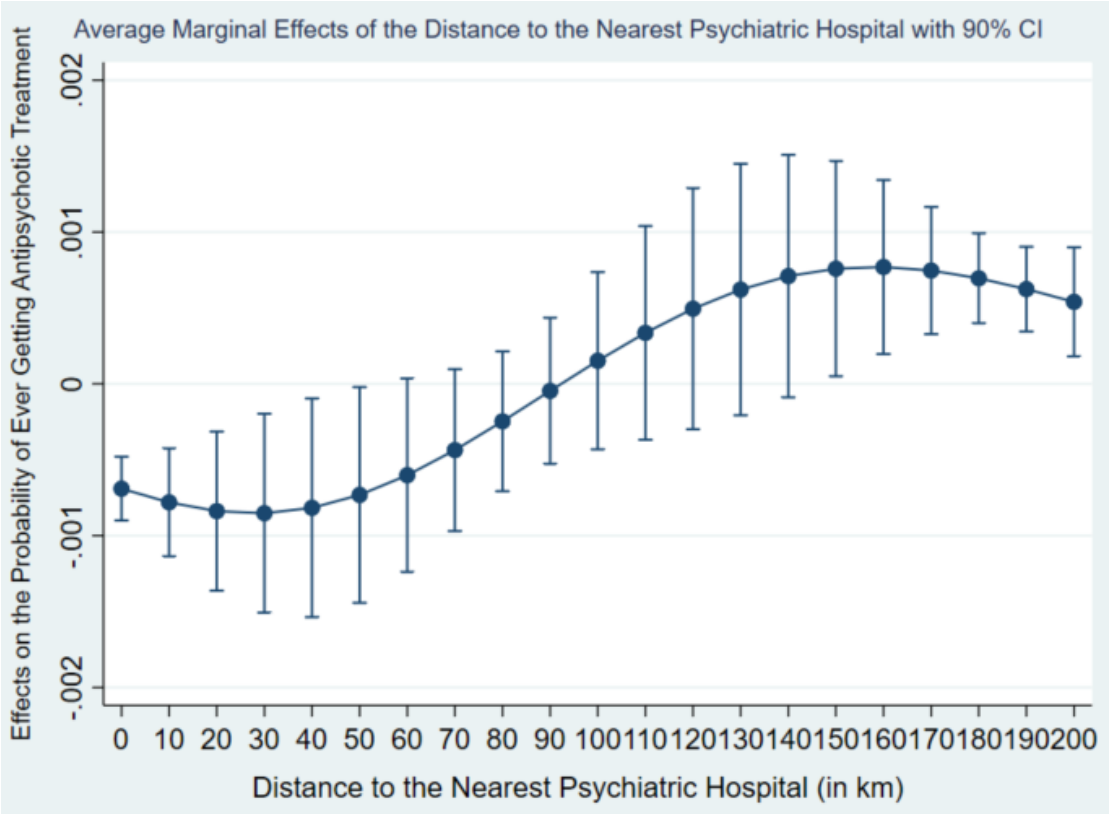

Figure 2.3 Average effect of distance to the nearest psychiatric service on the probability of ever receiving psychiatric service at different distance to the hospital (0-200 km)

Figures 2.4-2.6 show the relationship of patients' years of schooling to the probability

of any lifetime use of antipsychotic medication in communities with different proportions of minority residents (Figure 2.4), different proportions of rural residents (Figure 2.5), and at different distances from the nearest psychiatric hospital (Figure 2.6), after adjusting for all other variables in the model. Higher levels of education in the patient are consistently associated with a greater probability of lifetime use of antipsychotic medication, but this effect gets weaker as the proportion of ethnic minority residents in the village/neighborhood increases (becoming non-significant when the proportion exceeds 90%), gets stronger as the proportion of rural residents in the village/neighborhood gets greater (becoming statistically significant when the proportion exceeds 50%), and is statistically significant in villages/neighborhoods at intermediate distances from the nearest psychiatric hospital (30-130 km) but non-significant in villages/neighborhoods at the shortest (<30km) or longest (>130km) distances from the nearest psychiatric hospital.

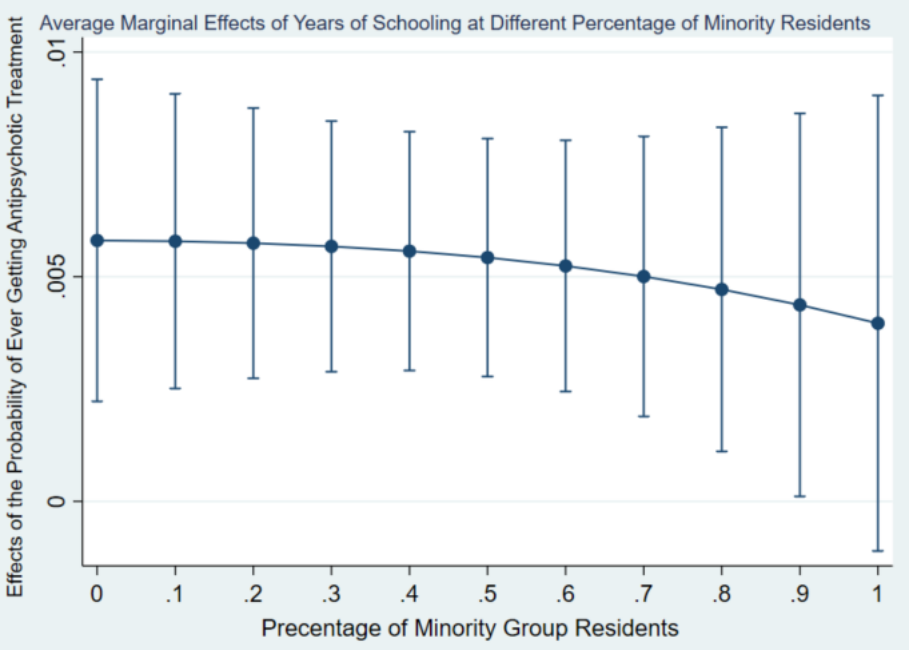

Figure 2.4. Effect of years of schooling on the probability of getting medication at different percentage of Muslim residents with 90% confidence interval

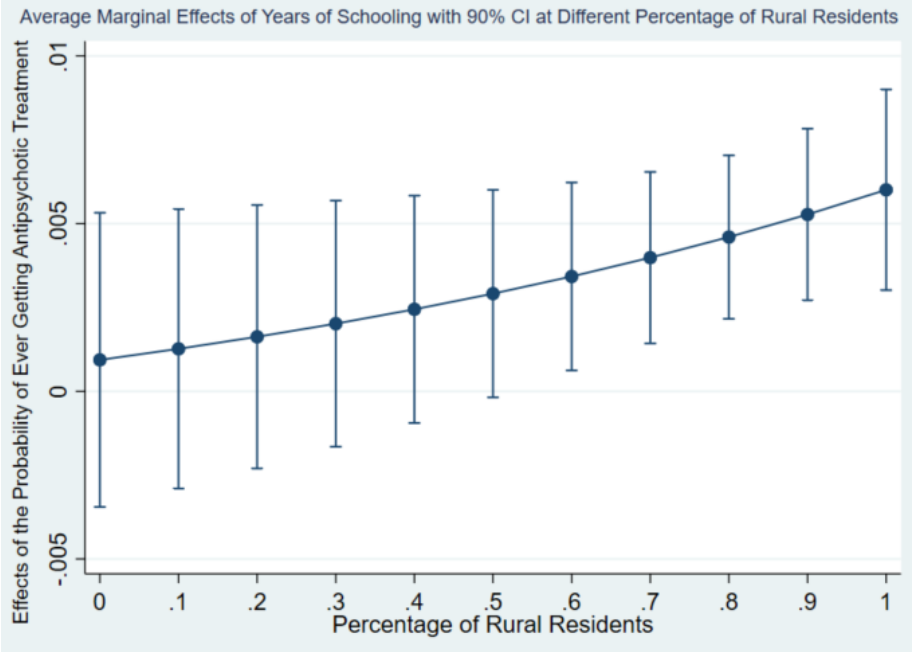

Figure 2.5. Effect of years of schooling on the probability of getting medication at different percentage of rural residents with 90% confidence interval

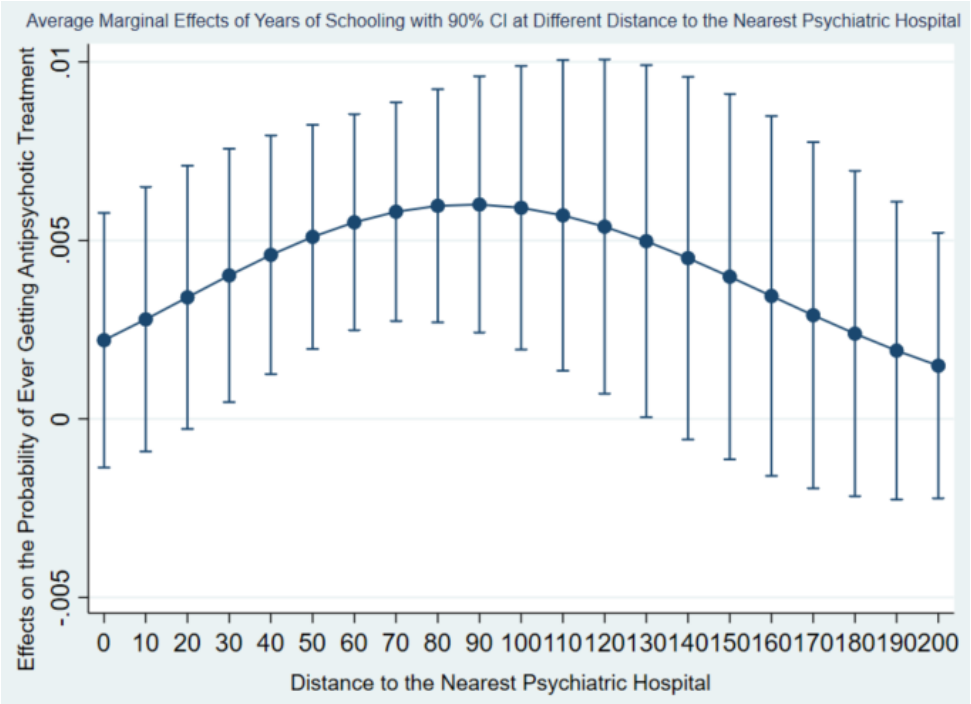

Figure 2.6. Effect of years of schooling on the probability of getting medication at different distance to the nearest psychiatric service

Figures 2.7-2.9 show the relationship of patients' age to the probability of any lifetime use of antipsychotic medication in communities with different proportions of

minority residents (Figure 2.7), different proportions of rural residents (Figure 2.8), and at different distances from the nearest psychiatric hospital (Figure 2.9), after adjusting for all other variables in the model. Greater patient age at the time of the survey is associated with a decreased probability of lifetime use of antipsychotic medication, but this effect gets stronger as the proportion of ethnic minority residents in the village/neighborhood increases (becoming statistically significant at the $p<0.10$ level when the proportion of ethnic minority residents exceeds 30%), gets weaker as the proportion of rural residents in the village/neighborhood increases (remaining statistically significant even for locations with 100% rural residents), and gets stronger as the distance of the village/neighborhood to the nearest psychiatric hospitals increases up to about 140 km and then gradually gets weaker at distances to the nearest psychiatric hospital become greater than 140 km (become statistically significant in the 60-160 km range).

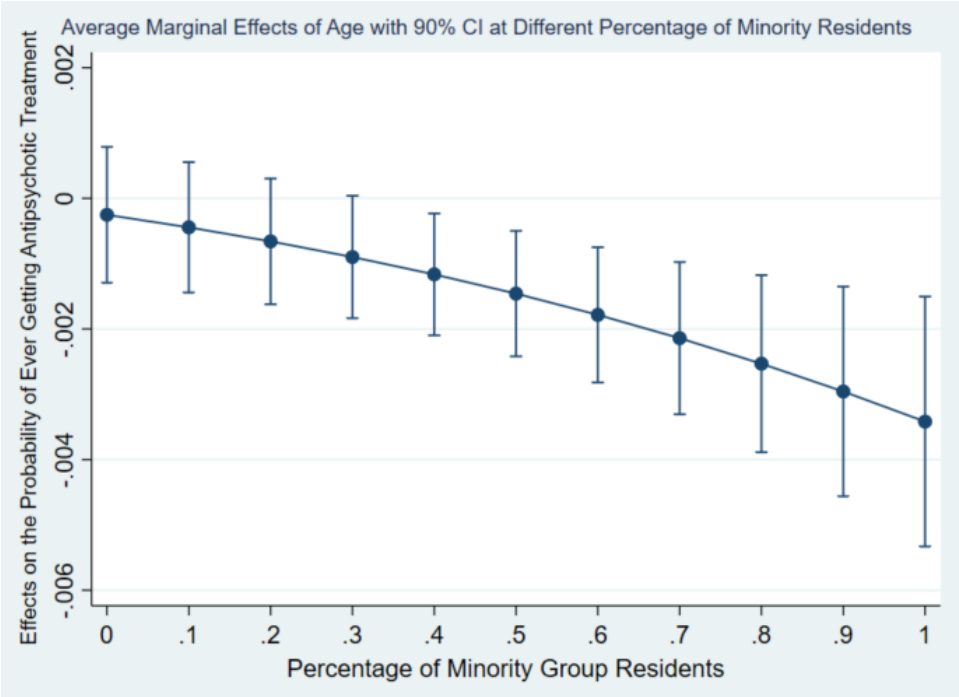

Figure 2.7 Effect of age on the probability of getting medication at different percentage of Muslim residents with 90% confidence interval

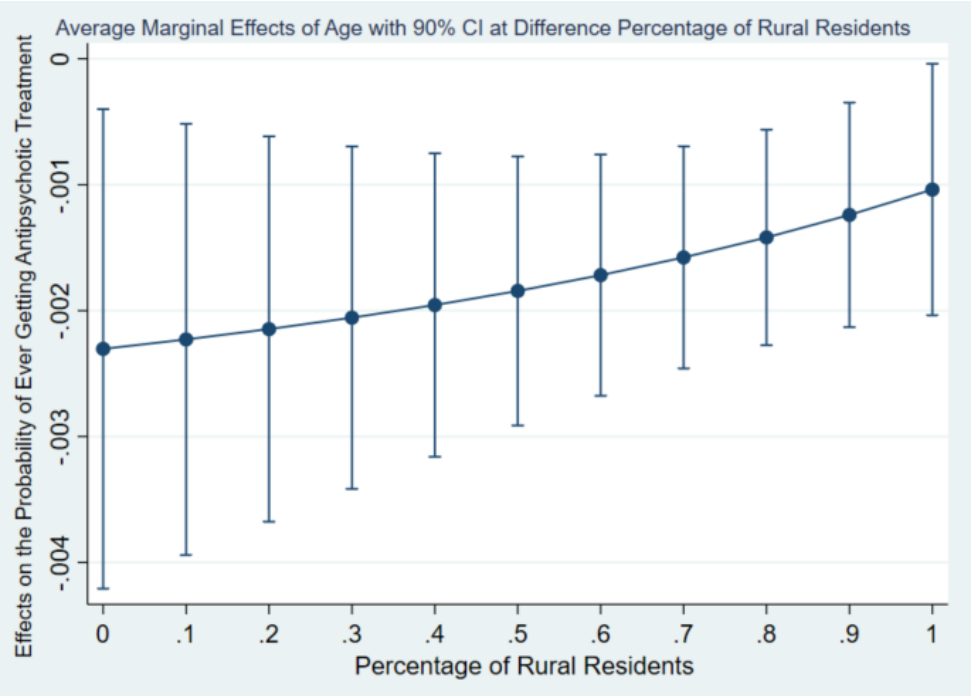

Figure 2.8 Effect of age on the probability of getting medication at different percentage of rural residents with 90% confidence interval

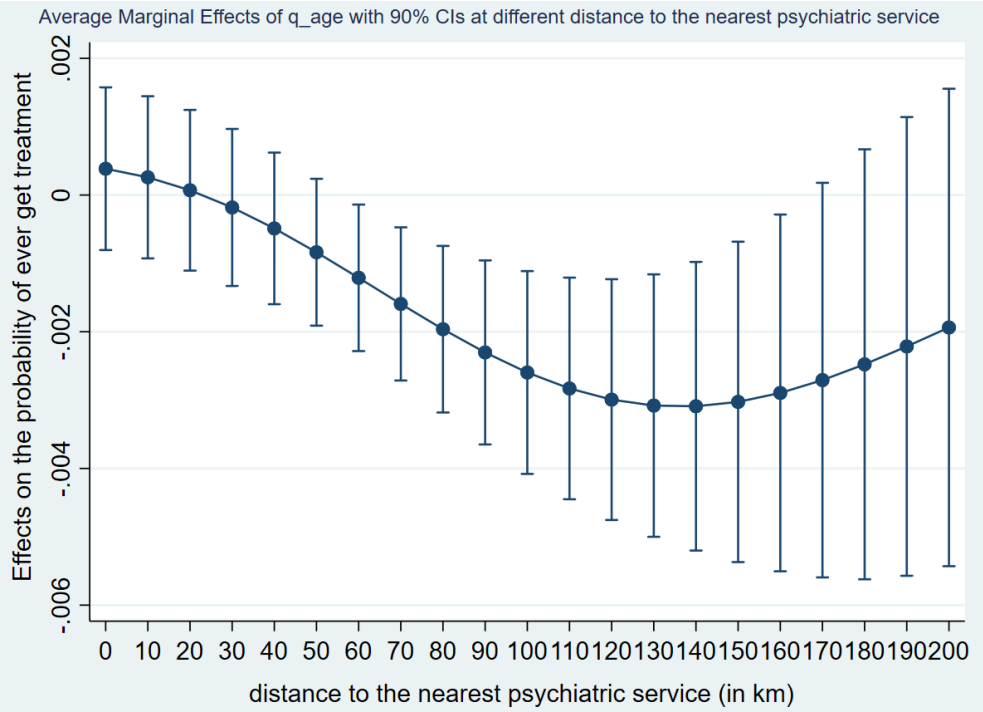

Figure 2.9 Effect of age on the probability of getting medication at different distance to the nearest psychiatric service

2.3.3 Any use of antipsychotic medication at any time in the prior three months among persons who had used antipsychotic medication at any time their lifetime

Among all the 2128 patients who ever used antipsychotic medication, 1683 (79.1%) had used antipsychotic medication at least once in the last three months. Results of the multilevel regression (equation 3) analysis to assess the probability using antipsychotic medication at any time in the prior three months is shown in Table 2.3.

Table 2.3. Multi-level analysis of any use antipsychotic medication in the three months prior to the survey (n=1891)[a]

	coefficient	OR (90% CI)	p
Individual-level characteristics			
gender (1=female 2=male)	**-0.307**	**0.736 (0.588-0.920)**	**0.024***
age in years	**-0.062**	**0.940 (0.918-0.962)**	**<0.001*****
ethnicity (1=Han 2=ethnic minority)	-0.185	0.831 (0.594-1.162)	0.364
Patient's household is classified 'rural household' (1=rural, 2=urban)	**0.513**	**1.671 (1.073-2.602)**	**0.057***
years of formal schooling	0.078	1.081 (0.933-1.253)	0.385
marital status (1=currently married, 2=other)	-0.081	0.923 (0.602-1.414)	0.756
patient's household classified as 'impoverished' (1=yes, 2=no)	**-0.410**	**0.664 (0.386-1.568)**	**0.053***
calendar year of initial onset of psychotic symptoms	-0.002	0.998 (0.987-1.009)	0.718
Ever hospitalized in psychiatric department (1=yes 2=no)	**-1.002**	**0.367 (0.291-0.464)**	**<0.001**
Followed up by primary care providers in the recent year (1=yes 2=no)	**-0.312**	**0.732 (0.585-0.915)**	**0.021***
Community-level characteristics			
percent rural residents	1.036	2.817 (0.506-15.675)	0.321
percent ethnic minority residents	**-1.626**	**0.197 (0.048-0.801)**	**0.057***
distance to the nearest psychiatric hospital (km)	-0.020	0.980 (0.958-1.002)	0.143
quadratic term of distance to the nearest psychiatric hospital (km)	-0.001	1.000 (1.000-1.000)	0.337
average GDP per capita of the county at time of survey	0.041	1.042 (0.991-1.059)	0.197
community is part of a county officially classified as 'mountainous'	0.120	1.128 (0.844-1.506)	0.495
has psychiatric hospital within the county (1=yes 2=no)	-0.162	0.850 (0.602-1.202)	0.441
Interaction terms			
years of schooling*percent rural residents in community	-0.050	0.952 (0.881-1.027)	0.287
years of schooling*percent ethnic minority residents in community	**0.087**	**1.091 (1.018-1.171)**	**0.040***
years of schooling*distance to the nearest psychiatric service	**0.001**	**1.001 (1.000-1.002)**	**0.030***
age*percent rural residents in community	**-0.030**	**0.971 (0.945-0.997)**	**0.068***
age*percent ethnic minority residents in community	**0.025**	**1.025 (1.002-1.049)**	**0.078***
age*distance to the nearest psychiatric service	**0.001**	**1.001 (1.000-1.001)**	**0.001****

[a] 251 cases were removed from the analysis because they have missing values in one or more covariates

Rho (residual intraclass correlation) = 0.042; Sigma_u =0.382 (90% CI: 0.157-0.926)

* p<0.1, ** p<0.01, *** p<0.001

Younger age, female gender, classified as urban household, prior psychiatric hospitalization, at least one primary care follow-up in the preceding year, registered as impoverished family and a higher proportion of Han ethnicity residents in the village/neighborhood were all associated with an increased likelihood of having used antipsychotic medication at any time in the prior three months.

There were several interactive effects of interest between the characteristics of patients and the characteristics of locations. Figures 2.10-2.12 show the relationship of patients' years of schooling and Figures 2.13-2.15 show the relationship of patients' age at the time of the survey to the probability of any use of antipsychotic medication over the prior three months in communities with different proportions of minority residents (Figures 2.10, 2.13), different proportions of rural residents (Figures 2.11, 2.14), and at different distances from the nearest psychiatric hospital (Figure 2.12, 2.15), after adjusting for all other variables in the model. The relationship of education as a predictor of any use of antipsychotic medication in the prior three months was negative and non-significant in locations where most residents were of Han ethnicity, but the relationship became a statistically significant positive relationship in communities where over 60% of the residents were from ethnic minorities (Figure 2.10). Higher educational level was a stronger positive predictor (but not statistically significant) in the most urbanized communities; this relationship gradually decreased in strength as the proportion of rural residents in the community increased (Figure 2.11). There was a non-significant negative relationship of patients' educational level and prediction of medication use in the prior three months among patients living in communities less than 50km for the nearest psychiatric hospital, but the relationship strengthened as the distance increased; education became a significant positive predictor of medication use in communities situated more than 80km from the nearest psychiatric hospital.

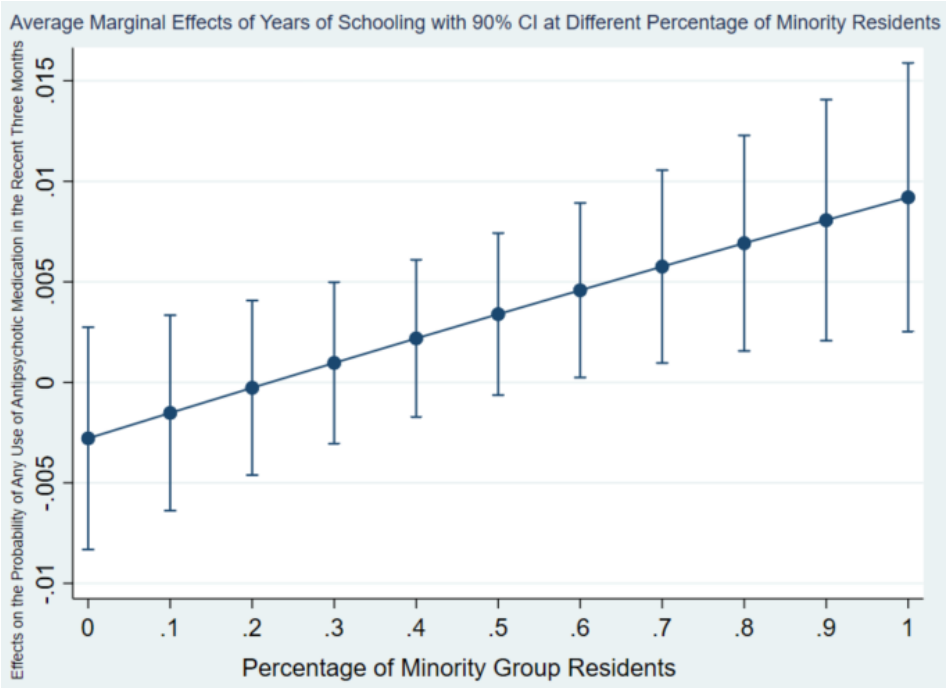

Figure 2.10 Effect of years of schooling on the probability of currently using antipsychotic medication at different percentage of Muslim residents with 90% confidence interval

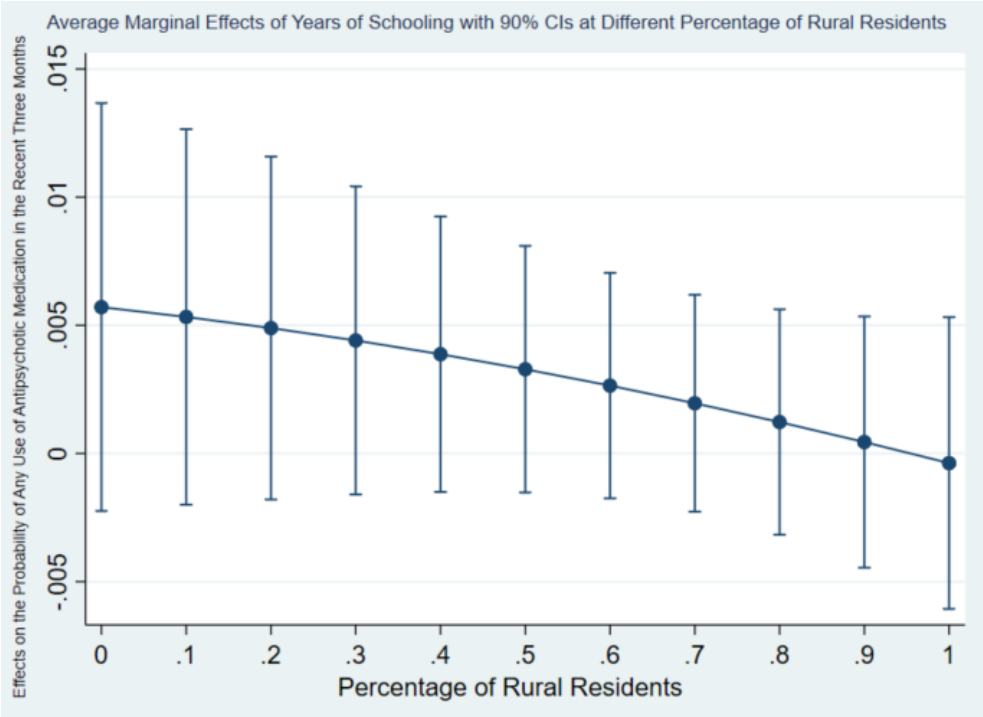

Figure 2.11 Effect of years of schooling on the probability of currently using antipsychotic medication at different percentage of rural residents with 90% confidence interval

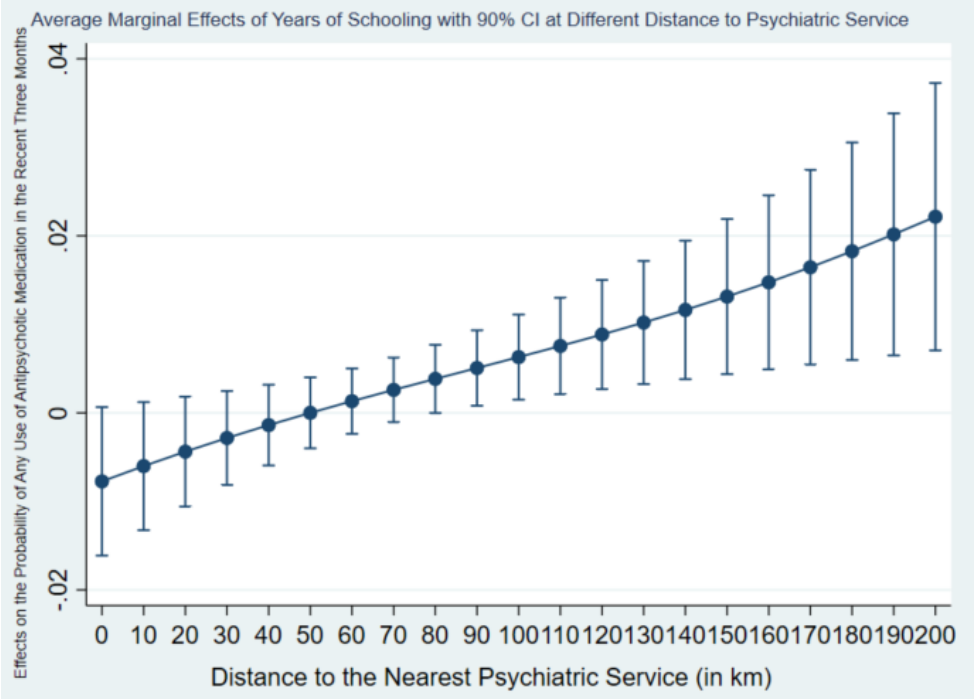

Figure 2.12 Effect of years of schooling on the probability of currently using antipsychotic medication at different distance to the nearest psychiatric service

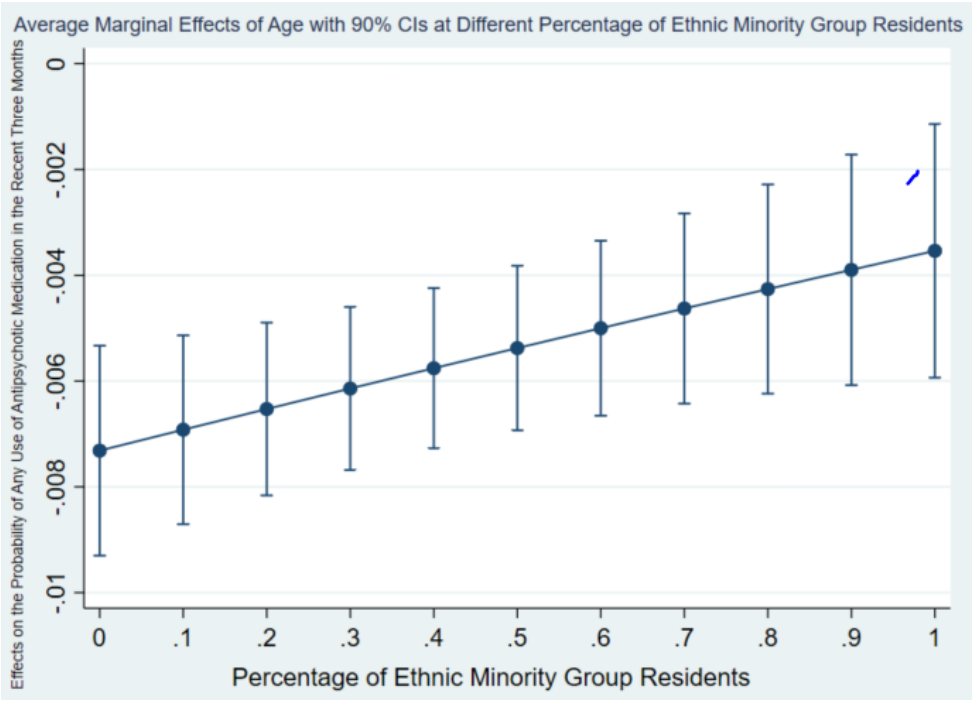

Figure 2.13 Effect of years of age the probability of currently using antipsychotic medication at different percentage of Muslim residents with 90% confidence interval

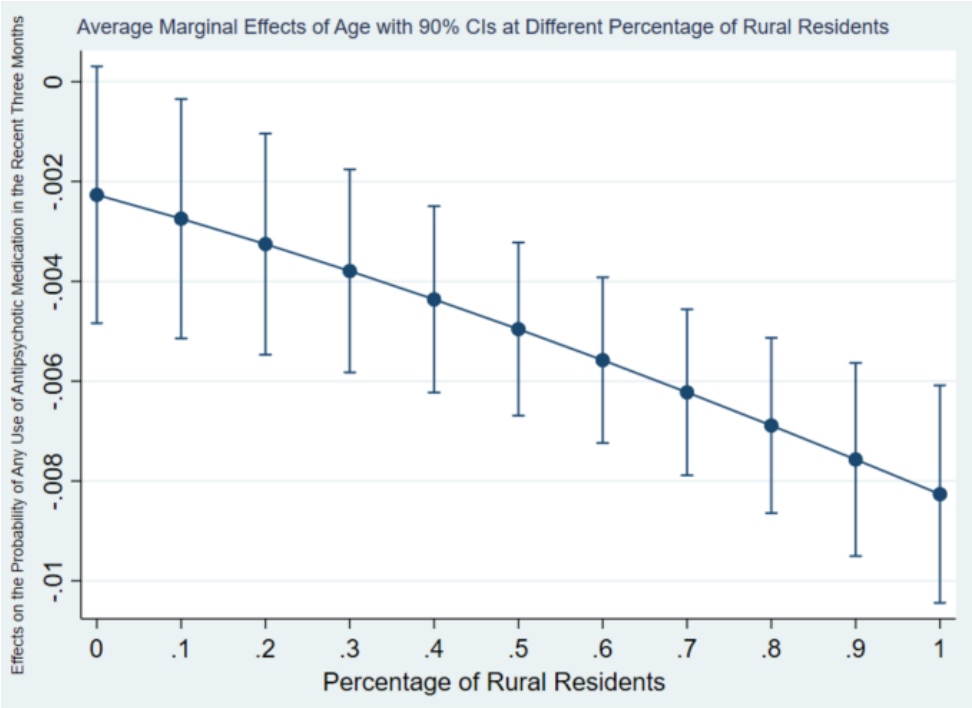

Figure 2.14 Effect of age on the probability of currently using antipsychotic medication at different percentage of rural residents with 90% confidence interval

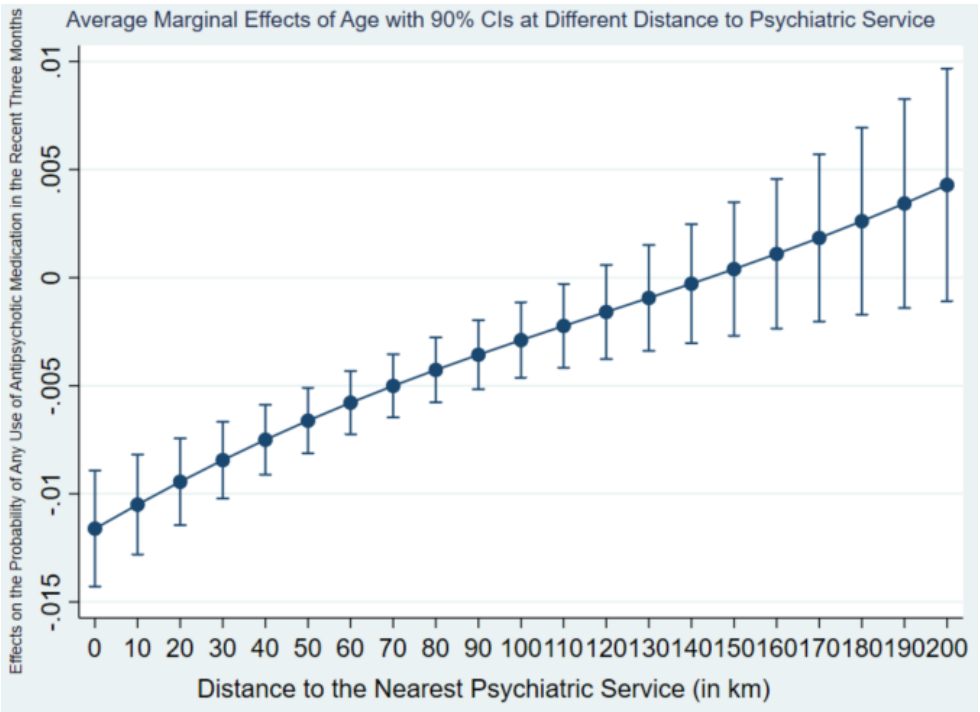

Figure 2.15 Effect of age on the probability of currently using antipsychotic medication at different distance to the nearest psychiatric service

Increasing age was associated with a decreased probability that the patient had taken any antipsychotic medication in the prior three months, a relationship that was statistically significant in communities regardless of the proportion of rural residents or the proportion of ethnic minority residents. However, this negative relationship weakened as the proportion of ethnic minority residents in the community increased (Figure 2.13) and strengthened as the proportion of rural residents in the community increased (Figure 2.14). Increasing age was a significant negative predictor of any antipsychotic medication use in the prior three months for patients living in villages/neighborhoods less than 100 km from the nearest psychiatric hospital, but age was a non-significant positive predictor of medication use for patients living in villages/neighborhoods more than 120 km from the nearest psychiatric hospital.

2.3.4 Regular medication use over the prior three months among patient who have ever used antipsychotic medication in the prior three months

Among all 1684 patients who had ever used antipsychotic medication in the prior three months, 1446 (85.9%) had regularly used the medication as prescribed.

Table 2.4 Multi-level analysis of regularly using anti-psychotic medication in the most recent three months (n=1,457)[a]

	coefficient	OR (90% CI)	p
Individual-level characteristics			
gender (1=female 2=male)	**0.395**	**1.485 (1.102-2.000)**	**0.029***
age in years	-0.008	0.992 (0.970-1.016)	0.590
ethnicity (1=Han 2=ethnic minority)	-0.021	0.980 (0.627-1.530)	0.940
Patient's household is classified 'rural household' (1=rural, 2=urban)	0.448	1.565 (0.952-2.573)	0.138
years of formal schooling	-0.023	0.977 (0.942-1.013)	0.287
marital status (1=currently married, 2=other)	0.141	1.151 (0.837-1.583)	0.467
patient's household classified as 'impoverished' (1=yes, 2=no)	-0.083	0.920 (0.594-1.425)	0.755
calendar year of initial onset of psychotic symptoms	-0.007	0.993 (0.978-1.009)	0.474
Ever hospitalized in psychiatric department (1=yes 2=no)	**-0.799**	**0.450 (0.329-0.615)**	**<0.001*** **
Followed up by primary care providers in the recent year (1=yes 2=no)	0.197	1.218 (0.887-1.671)	0.306
Receiving free antipsychotic medication from NISPP system (1=yes, 2=no)	**-0.311**	**0.733 (0.548-0.981)**	**0.079***
Receiving free antipsychotic medication from DPF system (1=yes, 2=no)	-0.015	0.985 (0.707-1.372)	0.940

Community-level characteristics			
percent rural residents	1.347	3.845 (0.969-15.259)	0.108
percent ethnic minority residents	0.166	1.180 (0.674-2.065)	0.626
distance to the nearest psychiatric hospital (km)	-0.003	0.998 (0.992-1.003)	0.441
average GDP per capita of the county at time of survey	-0.046	0.955 (0.907-1.004)	0.133
community is part of a county officially classified as 'mountainous'	0.027	1.028 (0.700-1.509)	0.907
has psychiatric hospital within the county (1=yes 2=no)	-0.332	0.717 (0.493-1.045)	0.146
Interaction terms			
age*percent rural residents in community	**-0.027**	**0.973 (0.949-0.999)**	**0.086***

[a] 227 cases were removed from the analysis because they have missing values in one or more covariates
Rho (residual intraclass correlation) <0.001; Sigma_u =0.003 (95% CI: 1.89e-9-3704.01)
* p<0.1, ** p<0.01, *** p<0.001

As shown in Table 2.4, gender is the only significant demographic variable related to the regular use of antipsychotic medication in the prior three months, and males are likely to maintain good adherence of drug use than female patients (OR=1.485, p=0.029). Moreover, two other patient-level variables – any prior psychiatric hospitalization and current receipt of free antipsychotic medication from NISPP – were significantly associated with regularly using antipsychotic medication in the prior three months. Interestingly, the receipt of free medication from the Disabled Persons' Federation was NOT a predictor of the regular use of medication; this may be due to the fact that the NISPP system provides both free medication AND regular follow-up by community-level health providers while the Disabled Persons' Federation only provides medication (without any follow-up monitoring).

None of the village/neighborhood-level was significantly associated with regular use of medication in the prior three months. As shown in Figure 2.16, after adjusting for other covariates, younger patients were more likely to regularly use medications in the prior three months, but this relationship was not statistically significant in predominantly urban communities—it was only significant in villages/neighborhoods in which more than 40% of residents were members of non-Han ethnic minorities.

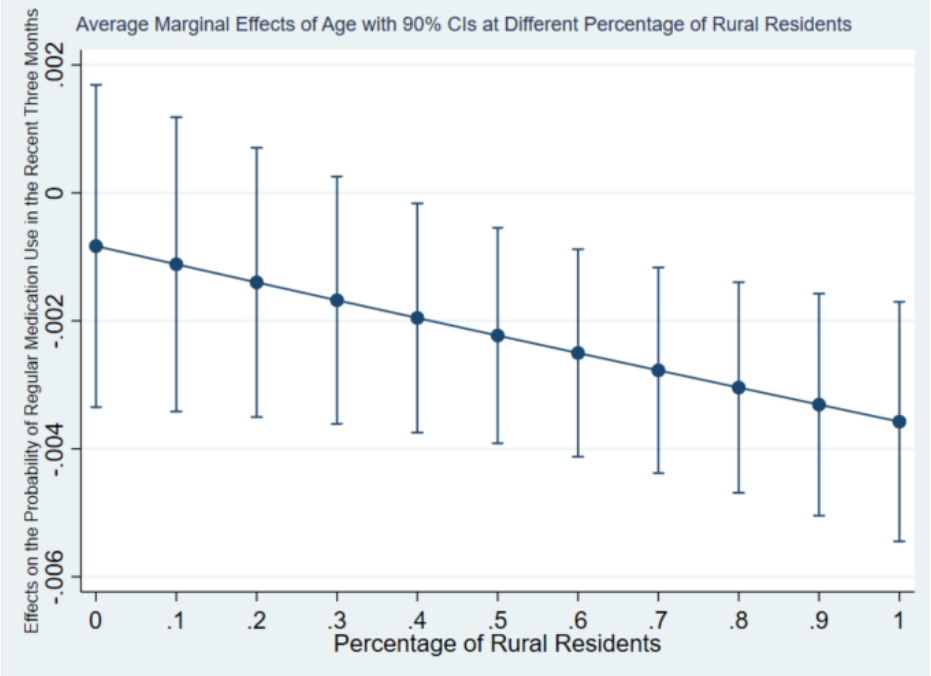

Figure 2.16 Effect of age on the probability of regularly using antipsychotic medication in the most recent three months at different percentage of rural residents with 90% confidence interval

2.4 Prediction

After adjusting for a variety of patient-level factors and community-level factors and for potential interactive effects of these two levels of factors, we estimate the probability that community-dwelling patients with schizophrenia have 1) ever used antipsychotic medication in their lifetime, 2) ever used antipsychotic medication in the prior three months, and 3) regularly used antipsychotic medications over the prior three months.

2.4.1 The predicted probability of ever using antipsychotics

The mean of predicted probability of patients with schizophrenia living in these villages/neighborhoods is 94.2% (SD 5.3%). As shown in Figure 2.17, the predicted probability of ever receiving in the two counties with more ethnic minority residents and in the mountainous area in the southwest of the province is much lower than the average level in the eight counties after controlling for all the individual-level factors included in the equation one.

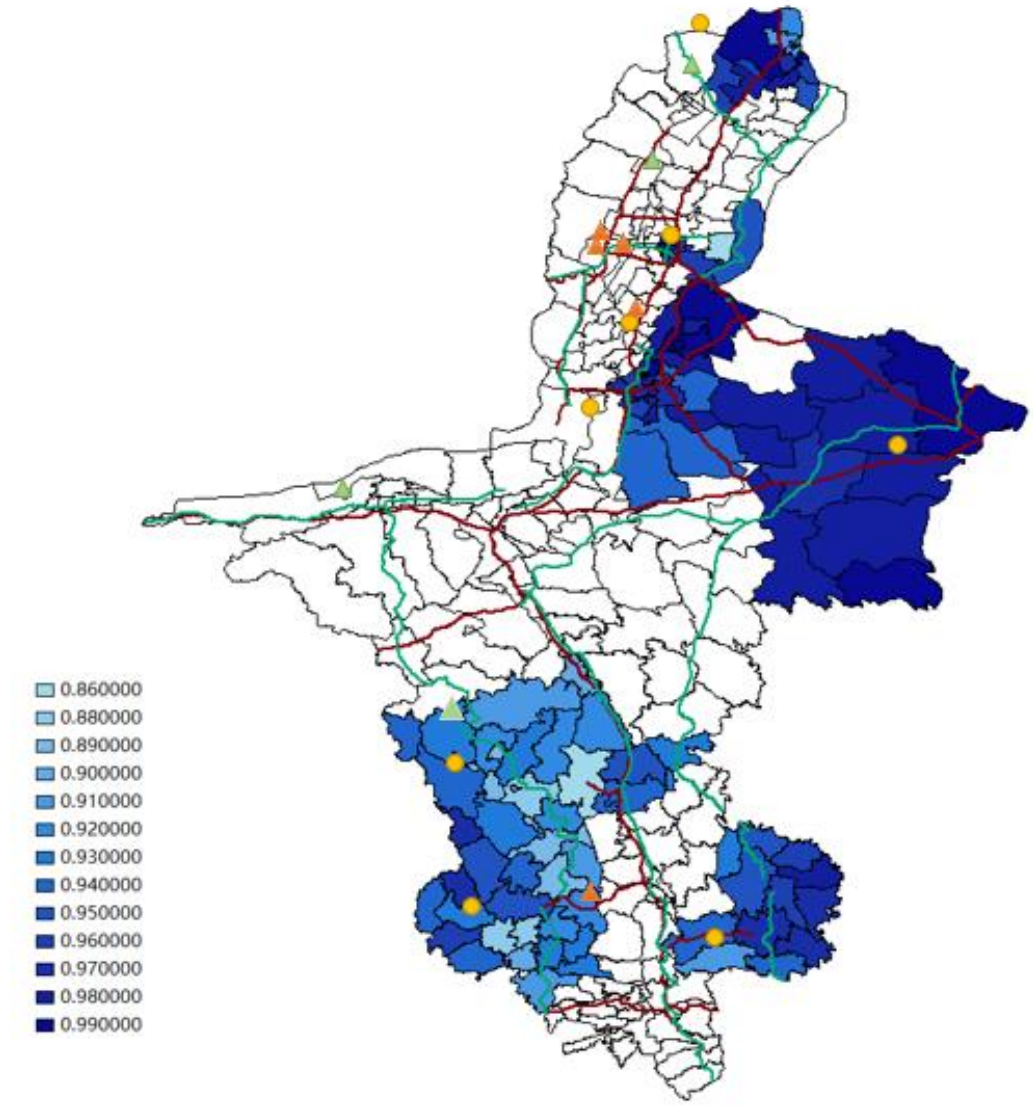

Figure 2.17 Average predicted probability of patients with schizophrenia ever receiving treatment in each township after controlling for individual-level factors (including age, gender, ethnicity, years of schooling, marital status, and years of disease onset)

Center of the county

Public psychiatric hospital

Private psychiatric hospital

2.4.2 The Predicted probability of any medication use in the three months prior to the survey

The mean of predicted probability of any use of antipsychotic among all the patients who ever used medication in the three months prior to the survey in these eight

counties is 72.2%. As shown in Figure 2.18, The predicted probability is highest in the urban area of the economic developed counties (Xingqing, Huinong, Lingwu, and Litong), followed by Yanchi where most residents are Han and not in the mountainous region. To our surprise, the predicted probability is lowest the rural area of two economic developed counties (Litong and Lingwu) where most residents are ethnic minority.

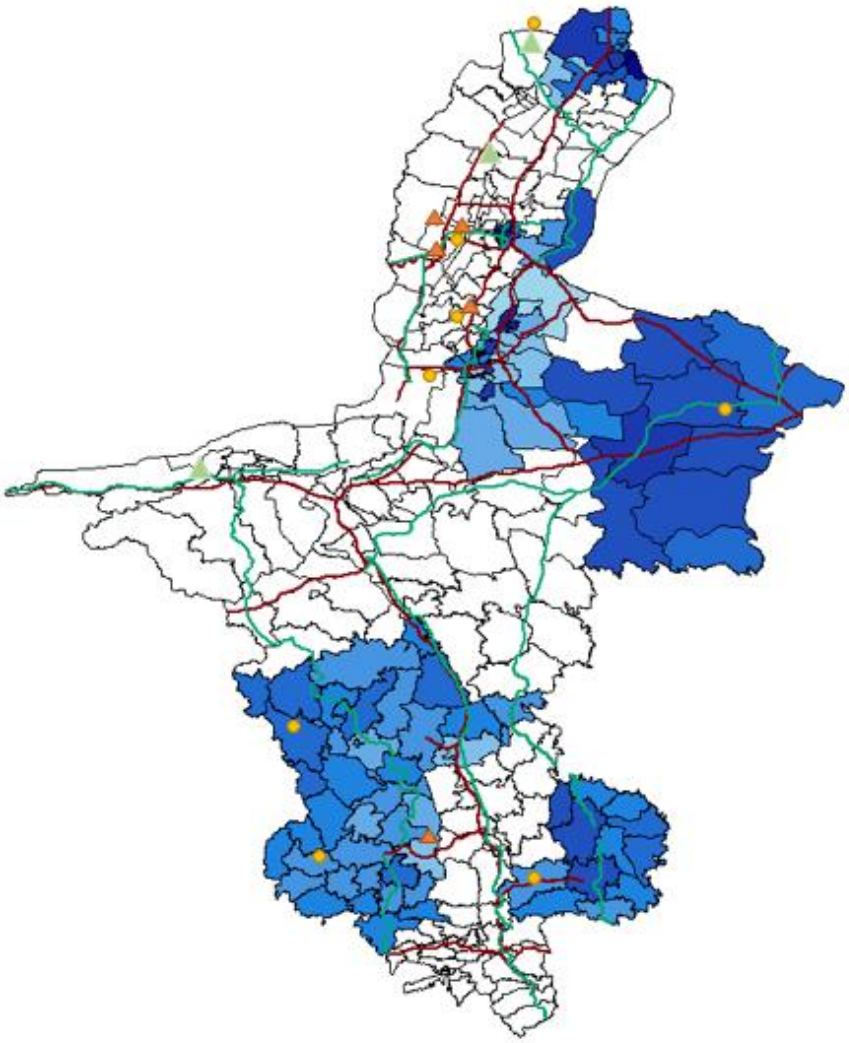

Figure 2.18 Average predicted probability of patients with schizophrenia using anti-psychotic medications at any time in the prior 3 months among all the patients ever using medication in each township after controlling for individual-level factors (including age, gender, ethnicity, years of schooling, marital status, age of onset, and psychiatric hospitalization history)

- Center of the county
- Public psychiatric hospital
- Private psychiatric hospital

2.4.3 The Predicted probability of regularly using medication use in the three months prior to the survey

The mean of predicted probability of regularly using antipsychotic medication in the three months prior to the survey is 83.6% among all the patients who ever used medication in these three months. Figure 2.19 indicates that the predicted probability is highest in the urban area of two counties where most residents are ethnic Han (Xingqing and Huinong), around 90%, while the predicted probability is just over 80% in the rural area of the rest of counties.

2.5 Robustness

2.5.1 random effect vs. fixed effects

We use random-effect model to estimate the cluster specific effect in equation 1, 2 and 3, which assumes the cluster effects follow normal distributions. Though fixed effect model does not make any assumption about the distribution of cluster effects, it is less efficient because it largely increases the number of coefficients to be estimated. In this study, we try to fit the corresponding equations with fixed cluster effects terms, but the models are unable to be fit due to the limitation of sample size and data separation (i.e., within single cluster, there is no variance of outcome variable). Therefore, we use random-effect models to estimate the coefficients in equation 1, 2, and 3.

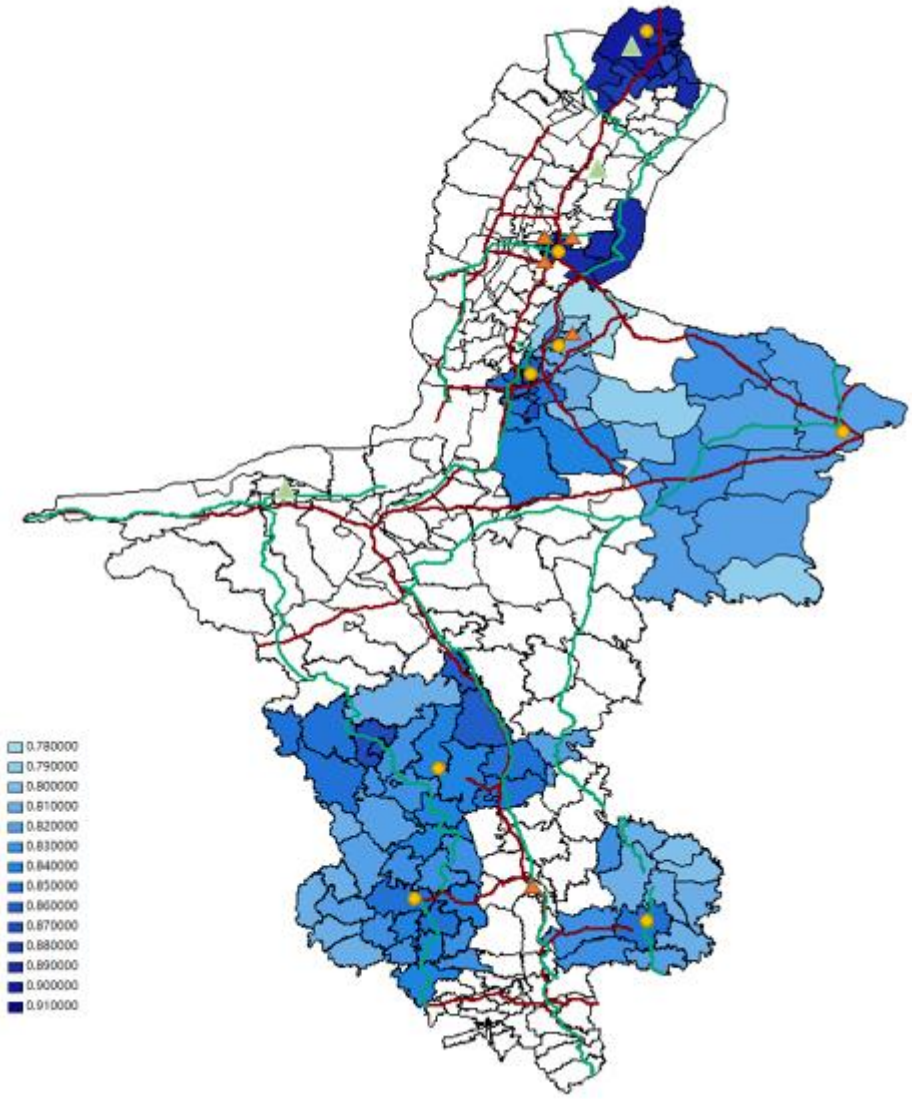

Figure 2.19 Average predicted probability of patients with schizophrenia using anti-psychotic medications regularly among all the patients using medication at any time in the prior 3 months in each township after controlling for individual-level factors (including age, gender, ethnicity, years of schooling, marital status, age of onset, and psychiatric hospitalization history)

- Center of the county
- Public psychiatric hospital
- Private psychiatric hospital

2.5.2 Year trend vs. year dummy

In equation one, two and three, we include a linear term to estimate the effect of the year of psychotic symptom onset. We have particular interest about the effect of year of psychotic symptom onset because it could control the unobservable effect of mental

health literacy and knowledge which have improve rapidly in China in these years. Here, we use fixed-effect year dummy to check the assumed linear relationship in these equations. As shown in the supplement materials (table 2.5-2.7), besides some minor difference (the quadratic term of distance to the nearest psychiatric service is a significant predictor in linear trend model to predict the probability of ever initiating treatment but not significant in year dummy model; average GDP of the county is not significant in the year trend model to predict either the probability of using antipsychotics anytime or regularly using medication in the three months before the survey but significant in the year dummy model; the interaction term between age and percentage of ethnic minority residents is a significant predictor of the probability of any use of medication in the three months prior to the survey but not significant in year dummy model; the percentage of rural residents is not significant to predict the probability of regularly using medication in the three months prior to the survey in the year trend model but significant in the year dummy model), the overall results remain consistent with that derived from equation 1 to 3.

Table 2.5. Multi-level analysis of ever using anti-psychotic medication using year of disease onset as fixed effect dummy variable (n=1,912)[a]

	coefficient	OR (90% CI)	p
Individual-level characteristics			
gender (1=female 2=male)	**-0.780**	**0.458 (0.230-0.705)**	**0.003****
age in years	-0.010	0.990 (0.935-1.049)	0.781
ethnicity (1=Han 2=ethnic minority)	-0.398	0.672 (0.358-1.259)	0.725
Patient's household is classified 'rural household' (1=rural, 2=urban)	0.342	1.408 (0.522-3.799)	0.570
years of formal schooling	0.055	1.056 (0.900-1.239)	0.574
marital status (1=currently married, 2=other)	-0.099	0.906 (0.570-1.440)	0.725
patient's household classified as 'impoverished' (1=yes, 2=no)	-0.001	0.999 (0.462-2.162)	0.999
Community-level characteristics			
percent rural residents	**-4.905**	**0.007 (0.001-0.458)**	**0.050***
percent ethnic minority residents	1.891	6.627 (0.570-77.028)	0.205
distance to the nearest psychiatric hospital (km)	0.004	1.004 (0.959-1.051)	0.881
quadratic term of distance to the nearest psychiatric hospital (km)	0.0001	1.000 (1.000-1.0003)	0.135
average GDP per capita of the county at time of survey	0.103	1.110 (0.984-1.250)	0.156
community is part of a county officially classified as 'mountainous'	0.072	1.074 (0.628-1.837)	0.826
has psychiatric hospital within the county (1=yes 2=no)	0.620	1.860 (0.816-4.240)	0.215
Interaction terms			
years of schooling*percent rural residents in community	0.148	1.160 (0.974-1.381)	0.163
years of schooling*percent ethnic minority residents in community	-0.096	0.909 (0.798-1.035)	0.228

years of schooling*distance to the nearest psychiatric service	-0.0003	1.000 (0.998-1.001)	0.746
age*percent rural residents in community	0.060	1.062 (0.998-1.130)	0.110
age*percent ethnic minority residents in community	**-0.048**	**0.953 (0.915-0.992)**	**0.049***
age*distance to the nearest psychiatric service	**-0.001**	**0.999 (0.999-1.000)**	**0.049***

[a] 251 cases were removed from the analysis because they have missing values in one or more covariates, and 117 cases were removed because data separation within one single year

Rho (residual intraclass correlation) = 0.308; Sigma_u =1.210 (95% CI: 0.653-1.705)

* p<0.1, ** p<0.01, *** p<0.001

Table 2.6. Multi-level analysis of currently using anti-psychotic medication using year of disease onset as fixed effect dummy variable (n=1,884)[a]

	coefficient	OR (90% CI)	p
Individual-level characteristics			
gender (1=female 2=male)	**-0.381**	**0.683 (0.539-0.920)**	**0.008****
age in years	**-0.062**	**0.940 (0.916-0.964)**	**<0.001*****
ethnicity (1=Han 2=ethnic minority)	-0.036	0.964 (0.674-1.380)	0.868
Patient's household is classified 'rural household' (1=rural, 2=urban)	**0.531**	**1.701 (1.073-2.695)**	**0.058***
years of formal schooling	-0.051	0.950 (0.884-1.021)	0.243
marital status (1=currently married, 2=other)	-0.142	0.867 (0.671-1.121)	0.362
patient's household classified as 'impoverished' (1=yes, 2=no)	**-0.429**	**0.664 (0.386-1.568)**	**0.053***
Ever hospitalized in psychiatric department (1=yes 2=no)	**-1.033**	**0.356 (0.278-0.456)**	**<0.001**
Followed up by primary care providers in the recent year (1=yes 2=no)	**-0.316**	**0.729 (0.576-0.923)**	**0.028***
Community-level characteristics			
percent rural residents	1.404	4.072 (0.662-25.059)	0.204
percent ethnic minority residents	**-1.725**	**0.178 (0.040-0.803)**	**0.059***
distance to the nearest psychiatric hospital (km)	-0.023	0.977 (0.954-1.001)	0.115
quadratic term of distance to the nearest psychiatric hospital (km)	-0.001	1.000 (1.000-1.000)	0.297
average GDP per capita of the county at time of survey	**0.056**	**1.058 (1.002-1.116)**	**0.086***
community is part of a county officially classified as 'mountainous'	0.080	1.083 (0.796-1.475)	0.670
has psychiatric hospital within the county (1=yes 2=no)	-0.109	0.896 (0.618-1.299)	0.628
Interaction terms			
years of schooling*percent rural residents in community	-0.073	0.929 (0.857-1.009)	0.142
years of schooling*percent ethnic minority residents in community	**0.080**	**1.083 (1.006-1.166)**	**0.077***
years of schooling*distance to the nearest psychiatric service	**0.001**	**1.001 (1.000-1.002)**	**0.008****
age*percent rural residents in community	**-0.035**	**0.966 (0.939-0.994)**	**0.044***
age*percent ethnic minority residents in community	0.024	1.024 (0.999-1.050)	0.115
age*distance to the nearest psychiatric service	**0.001**	**1.001 (1.000-1.001)**	**0.001****

[a] 251 cases were removed from the analysis because they have missing values in one or more covariates, and 7 cases were removed because data separation within one single year

Rho (residual intraclass correlation) = 0.064; Sigma_u =0.472 (90% CI: 0.157-0.926)

* p<0.1, ** p<0.01, *** p<0.001

Table 2.7. Multi-level analysis of regularly using anti-psychotic medication in the most recent three months using year of disease onset as fixed effect dummy variable (n=1396)[a]

	coefficient	OR (90% CI)	p
Individual-level characteristics			
gender (1=female 2=male)	**0.547**	**1.729 (1.259-2.374)**	**0.005****
age in years	-0.002	0.998 (0.973-1.023)	0.871
ethnicity (1=Han 2=ethnic minority)	-0.208	0.811 (0.504-1.305)	0.468
Patient's household is classified 'rural household' (1=rural, 2=urban)	0.449	1.566 (0.929-2.641)	0.158
years of formal schooling	-0.022	0.978 (0.942-1.016)	0.334
marital status (1=currently married, 2=other)	0.222	1.248 (0.894-1.743)	0.274
patient's household classified as 'impoverished' (1=yes, 2=no)	-0.198	0.920 (0.594-1.425)	0.470
Ever hospitalized in psychiatric department (1=yes 2=no)	**-0.777**	**0.460 (0.329-0.643)**	**<0.001*****
Followed up by primary care providers in the recent year (1=yes 2=no)	0.179	1.218 (0.887-1.671)	0.375
Receiving free antipsychotic medication from NISPP system (1=yes, 2=no)	**-0.348**	**0.706 (0.517-0.964)**	**0.066***
Receiving free antipsychotic medication from DPF system (1=yes, 2=no)	0.098	1.103 (0.777-1.566)	0.644
Community-level characteristics			
percent rural residents	**1.529**	**4.613 (1.048-20.298)**	**0.090***
percent ethnic minority residents	0.398	1.489 (0.815-2.719)	0.277
distance to the nearest psychiatric hospital (km)	-0.004	0.996 (0.991-1.002)	0.299
average GDP per capita of the county at time of survey	**-0.059**	**0.955 (0.907-1.004)**	**0.076***
community is part of a county officially classified as 'mountainous'	0.183	1.201 (0.800-1.803)	0.458
has psychiatric hospital within the county (1=yes 2=no)	-0.320	0.726 (0.491-1.073)	0.178
Interaction terms			
age*percent rural residents in community	**-0.029**	**0.971 (0.945-0.999)**	**0.085***

[a] 227 cases were removed from the analysis because they have missing values in one or more covariates, and 39 cases were removed because data separation within one single year

Rho (residual intraclass correlation) <0.001; Sigma_u =0.002 (95% CI: 4.48e-12-9.26e5)

* p<0.1, ** p<0.01, *** p<0.001

2.5.3 single-level model vs. multi-level model

To compare with the results of multi-level models, we also conduct the single-level analysis corresponding to equation 1-3 without the cluster specific term, ζ_j, η_j, or κ_j (shown in table 2.8-2.10). The results of single-level models and multilevel models are similar to (except for two minor differences: the percentage of rural residents is a significant predictor of the probability of ever initiating treatment in multilevel models but not significant in single-level estimation; ethnic minority is a significant predictor of the probability of ever initiating psychotic treatment in the multilevel model but not significant in the single level model). The coefficients in multilevel and single level to

estimate the probability of regularly medication in the three months prior to the survey are consistent, which indicates there is very little cluster effects exist in equation three (the rho in equation three is very close to zero also confirms the individuals are fairly independent when estimating equation three). The rho value in equation one and two are 0.253 and 0.042 (indicating 25.3% of the total variance in equation one could be explained by cluster variance and 4.2% of the total variance in equation two could be explained by cluster variance). As rule of thumb, rho greater than 0.1 indicates that multilevel should be used, so use multilevel models in this study.

Table 2.8. Single-level analysis of ever using antipsychotic medication (n=2028)[a]

	coefficient	OR (90% CI)	p
Individual-level characteristics			
gender (1=female 2=male)	**-0.671**	**0.511 (0.360-0.722)**	**0.002****
age in years	-0.009	0.992 (0.945-1.040)	0.770
ethnicity (1=Han 2=ethnic minority)	**-0.529**	**0.589 (0.355-0.978)**	**0.086***
Patient's household is classified 'rural household' (1=rural, 2=urban)	0.134	1.144 (0.534-2.453)	0.771
years of formal schooling	0.092	1.081 (0.933-1.253)	0.270
marital status (1=currently married, 2=other)	-0.029	0.971 (0.660-1.429)	0.901
patient's household classified as 'impoverished' (1=yes, 2=no)	-0.222	0.801 (0.423-1.514)	0.566
calendar year of initial onset of psychotic symptoms	**0.019**	**1.019 (1.003-1.036)**	**0.046***
Community-level characteristics			
percent rural residents	-3.291	0.037 (0.001-1.014)	0.101
percent ethnic minority residents	1.712	5.543 (0.736-41.712)	0.163
distance to the nearest psychiatric hospital (km)	-0.0004	1.000 (0.964-1.037)	0.983
quadratic term of distance to the nearest psychiatric hospital (km)	**0.0002**	**1.000 (1.000-1.0003)**	**0.086***
average GDP per capita of the county at time of survey	0.093	1.097 (0.993-1.213)	0.128
community is part of a county officially classified as 'mountainous'	0.041	1.042 (0.704-1.543)	0.863
has psychiatric hospital within the county (1=yes 2=no)	0.405	1.500 (0.800-2.810)	0.288
Interaction terms			
years of schooling*percent rural residents in community	0.073	1.075 (0.932-1.241)	0.404
years of schooling*percent ethnic minority residents in community	-0.092	0.912 (0.819-1.015)	0.158
years of schooling*distance to the nearest psychiatric service	-0.0002	1.000 (0.999-1.001)	0.863
age*percent rural residents in community	0.045	1.046 (0.995-1.010)	0.136
age*percent ethnic minority residents in community	**-0.039**	**0.962 (0.932-0.994)**	**0.049***
age*distance to the nearest psychiatric service	**-0.001**	**0.999 (0.999-1.000)**	**0.046***

[a] 252 cases were removed from the analysis because they have missing values in one or more covariates

* p<0.1, ** p<0.01, *** p<0.001

Table 2.9. Single-level analysis of currently using anti-psychotic medication (n=1891)[a]

	coefficient	OR (90% CI)	p
Individual-level characteristics			
gender (1=female 2=male)	**-0.291**	**0.747 (0.602-0.929)**	**0.027***
age in years	**-0.060**	**0.942 (0.921-0.963)**	**<0.001*****
ethnicity (1=Han 2=ethnic minority)	-0.179	0.836 (0.604-1.159)	0.366
Patient's household is classified 'rural household' (1=rural, 2=urban)	**0.566**	**1.671 (1.073-2.602)**	**0.025***
years of formal schooling	0.040	0.960 (0.900-1.025)	0.306
marital status (1=currently married, 2=other)	-0.081	0.884 (0.699-1.119)	0.756
patient's household classified as 'impoverished' (1=yes, 2=no)	**-0.378**	**0.685 (0.491-0.955)**	**0.061***
calendar year of initial onset of psychotic symptoms	-0.003	0.998 (0.987-1.009)	0.691
Ever hospitalized in psychiatric department (1=yes 2=no)	**-0.973**	**0.368 (0.303-0.472)**	**<0.001**
Followed up by primary care providers in the recent year (1=yes 2=no)	**-0.303**	**0.739 (0.595-0.918)**	**0.022***
Community-level characteristics			
percent rural residents	1.101	3.006 (0.563-16.053)	0.280
percent ethnic minority residents	**-1.534**	**0.215 (0.055-0.838)**	**0.063***
distance to the nearest psychiatric hospital (km)	-0.020	0.980 (0.958-1.002)	0.127
quadratic term of distance to the nearest psychiatric hospital (km)	>-0.001	1.000 (0.999-1.000)	0.392
average GDP per capita of the county at time of survey	0.037	1.037 (0.990-1.087)	0.196
community is part of a county officially classified as 'mountainous'	0.113	1.120 (0.844-1.506)	0.498
has psychiatric hospital within the county (1=yes 2=no)	-0.165	0.848 (0.615-1.169)	0.399
Interaction terms			
years of schooling*percent rural residents in community	-0.052	0.949 (0.880-1.023)	0.253
years of schooling*percent ethnic minority residents in community	**0.081**	**1.084 (1.014-1.160)**	**0.047***
years of schooling*distance to the nearest psychiatric service	**0.001**	**1.001 (1.000-1.002)**	**0.032***
age*percent rural residents in community	**-0.029**	**0.971 (0.946-0.997)**	**0.068***
age*percent ethnic minority residents in community	**0.024**	**1.024 (1.000-1.047)**	**0.083***
age*distance to the nearest psychiatric service	**0.001**	**1.001 (1.000-1.001)**	**0.001****

[a] 251 cases were removed from the analysis because they have missing values in one or more covariates

* p<0.1, ** p<0.01, *** p<0.001

Table 2.10. Single-level analysis of regularly using anti-psychotic medication in the most recent three months (n=1,457)[a]

	coefficient	OR (90% CI)	p
Individual-level characteristics			
gender (1=female 2=male)	**0.395**	**1.485 (1.102-2.000)**	**0.029***
age in years	-0.008	0.992 (0.970-1.016)	0.590
ethnicity (1=Han 2=ethnic minority)	-0.021	0.980 (0.627-1.530)	0.940
Patient's household is classified 'rural household' (1=rural, 2=urban)	0.448	1.565 (0.952-2.573)	0.138
years of formal schooling	-0.023	0.977 (0.942-1.013)	0.287
marital status (1=currently married, 2=other)	0.141	1.151 (0.837-1.583)	0.467
patient's household classified as 'impoverished' (1=yes, 2=no)	-0.083	0.920 (0.594-1.425)	0.755
calendar year of initial onset of psychotic symptoms	-0.007	0.993 (0.978-1.009)	0.474
Ever hospitalized in psychiatric department (1=yes 2=no)	**-0.799**	**0.450 (0.329-0.615)**	**<0.001*****
Followed up by primary care providers in the recent year (1=yes 2=no)	0.197	1.218 (0.887-1.671)	0.306

Receiving free antipsychotic medication from NISPP system (1=yes, 2=no)	**-0.311**	**0.733 (0.548-0.981)**	**0.079***
Receiving free antipsychotic medication from DPF system (1=yes, 2=no)	-0.015	0.985 (0.707-1.372)	0.940
Community-level characteristics			
percent rural residents	1.347	3.845 (0.969-15.259)	0.108
percent ethnic minority residents	0.166	1.180 (0.674-2.065)	0.626
distance to the nearest psychiatric hospital (km)	-0.003	0.998 (0.992-1.003)	0.441
average GDP per capita of the county at time of survey	-0.046	0.955 (0.907-1.004)	0.133
community is part of a county officially classified as 'mountainous'	0.027	1.028 (0.700-1.509)	0.907
has psychiatric hospital within the county (1=yes 2=no)	-0.332	0.717 (0.493-1.045)	0.146
Interaction terms			
age*percent rural residents in community	**-0.027**	**0.973 (0.949-0.999)**	**0.086***

[a] 227 cases were removed from the analysis because they have missing values in one or more covariates

Rho (residual intraclass correlation) <0.001; Sigma_u =0.003 (95% CI: 1.89e-9-3704.01)

* p<0.1, ** p<0.01, *** p<0.001

2.5.4 using township as the cluster

Equation 1-3 use village or community as the cluster in the models, and there are in total 990 cluster specific terms to be estimated (equals to the number of villages or communities in the dataset). Due to the concern that the sample size is not sufficient to analyze such large number of specific terms, we here use township, the higher administrative unit then village or community, as cluster to reduce the number of cluster specific term to 138. The results of corresponding multilevel models as equation 1-3 are shown in the appendix table 2.11-2.13 and consistent with those of multilevel models using villages or communities as the cluster (except for one minor difference: percentage of rural residents is a significant predictor of the probability of ever receiving treatment in the model using villages or communities as the cluster but not in the models using townships as clusters).

Table 2.11. Multi-level analysis of ever using antipsychotic medication using township as the cluster (n=2,028)[a]

	coefficient	OR (90% CI)	p
Individual-level characteristics			
gender (1=female 2=male)	**-0.693**	**0.500 (0.350-0.716)**	**0.001****
age in years	-0.010	0.990 (0.943-1.040)	0.742
ethnicity (1=Han 2=ethnic minority)	-0.510	0.600 (0.356-1.012)	0.108
Patient's household is classified 'rural household' (1=rural, 2=urban)	0.143	1.186 (0.537-2.622)	0.794
years of formal schooling	0.087	1.091 (0.950-1.253)	0.300

marital status (1=currently married, 2=other)	-0.011	0.989 (0.668-1.465)	0.963
patient's household classified as 'impoverished' (1=yes, 2=no)	-0.252	0.778 (0.386-1.568)	0.555
calendar year of initial onset of psychotic symptoms	**0.020**	**1.020 (1.004-1.037)**	**0.040***
Community-level characteristics			
percent rural residents	-3.322	0.036 (0.001-1.033)	0.103
percent ethnic minority residents	1.739	5.689 (0.724-44.698)	0.165
distance to the nearest psychiatric hospital (km)	-0.004	0.996 (0.959-1.035)	0.869
quadratic term of distance to the nearest psychiatric hospital (km)	**0.0002**	**1.000 (1.000-1.0003)**	**0.083***
average GDP per capita of the county at time of survey	0.084	1.087 (0.979-1.208)	0.189
community is part of a county officially classified as 'mountainous'	0.043	1.044 (0.680-1.603)	0.868
has psychiatric hospital within the county (1=yes 2=no)	0.437	1.673 (0.787-3.553)	0.285
Interaction terms			
years of schooling*percent rural residents in community	0.074	1.077 (0.932-1.244)	0.400
years of schooling*percent ethnic minority residents in community	-0.096	0.909 (0.814-1.014)	0.152
years of schooling*distance to the nearest psychiatric service	>-0.001	1.000 (0.999-1.001)	0.921
age*percent rural residents in community	0.046	1.047 (0.995-1.101)	0.138
age*percent ethnic minority residents in community	**-0.039**	**0.962 (0.931-0.994)**	**0.050***
age*distance to the nearest psychiatric service	**-0.0005**	**1.000 (0.999-1.000)**	**0.054***

[a] 251 cases were removed from the analysis because they have missing values in one or more covariates
Rho (residual intraclass correlation) = 0.042; Sigma_u =0.382 (95% CI: 0.113-1.286)
* p<0.1, ** p<0.01, *** p<0.001

Table 2.12. Multi-level analysis of currently using anti-psychotic medication using township as clusters (n=1,891)[a]

	coefficient	OR (90% CI)	p
Individual-level characteristics			
gender (1=female 2=male)	**-0.296**	**0.736 (0.588-0.920)**	**0.026***
age in years	**-0.061**	**0.940 (0.918-0.962)**	**<0.001*****
ethnicity (1=Han 2=ethnic minority)	-0.173	0.831 (0.594-1.162)	0.387
Patient's household is classified 'rural household' (1=rural, 2=urban)	**0.502**	**1.671 (1.073-2.602)**	**0.062***
years of formal schooling	0.045	1.081 (0.933-1.253)	0.264
marital status (1=currently married, 2=other)	-0.136	0.923 (0.602-1.414)	0.347
patient's household classified as 'impoverished' (1=yes, 2=no)	**-0.382**	**0.664 (0.386-1.568)**	**0.065***
calendar year of initial onset of psychotic symptoms	-0.002	0.998 (0.987-1.009)	0.731
Ever hospitalized in psychiatric department (1=yes 2=no)	**-0.980**	**0.367 (0.291-0.464)**	**<0.001**
Followed up by primary care providers in the recent year (1=yes 2=no)	**-0.306**	**0.732 (0.585-0.915)**	**0.021***
Community-level characteristics			
percent rural residents	0.959	2.817 (0.506-15.675)	0.358
percent ethnic minority residents	**-1.557**	**0.197 (0.048-0.801)**	**0.063***
distance to the nearest psychiatric hospital (km)	-0.021	0.980 (0.958-1.002)	0.118
quadratic term of distance to the nearest psychiatric hospital (km)	-0.0004	1.000 (1.000-1.000)	0.427
average GDP per capita of the county at time of survey	0.036	1.042 (0.991-1.059)	0.240
community is part of a county officially classified as 'mountainous'	0.093	1.128 (0.844-1.506)	0.602
has psychiatric hospital within the county (1=yes 2=no)	-0.170	0.850 (0.602-1.202)	0.414
Interaction terms			
years of schooling*percent rural residents in community	-0.048	0.952 (0.881-1.027)	0.300
years of schooling*percent ethnic minority residents in community	**0.084**	**1.091 (1.018-1.171)**	**0.043***
years of schooling*distance to the nearest psychiatric service	**0.001**	**1.001 (1.000-1.002)**	**0.031***

age*percent rural residents in community	**-0.028**	**0.971 (0.945-0.997)**	**0.080***
age*percent ethnic minority residents in community	**0.024**	**1.025 (1.002-1.049)**	**0.089***
age*distance to the nearest psychiatric service	**0.001**	**1.001 (1.000-1.001)**	**0.001****

[a] 251 cases were removed from the analysis because they have missing values in one or more covariates

Rho (residual intraclass correlation) = 0.015; Sigma_u =0.222 (90% CI: 0.070-0.710)

* p<0.1, ** p<0.01, *** p<0.001

Table 2.13. Multi-level analysis of regularly using anti-psychotic medication in the most recent three months (n=1,458)[a]

	coefficient	OR (90% CI)	p
Individual-level characteristics			
gender (1=female 2=male)	**0.395**	**1.485 (1.041-2.118)**	**0.029***
age in years	-0.008	0.992 (0.965-1.020)	0.590
ethnicity (1=Han 2=ethnic minority)	-0.020	0.980 (0.576-1.667)	0.940
Patient's household is classified 'rural household' (1=rural, 2=urban)	0.448	1.565 (0.866-2.829)	0.138
years of formal schooling	-0.023	0.977 (0.936-1.020)	0.287
marital status (1=currently married, 2=other)	0.141	1.151 (0.788-1.682)	0.467
patient's household classified as 'impoverished' (1=yes, 2=no)	-0.083	0.920 (0.546-1.550)	0.755
calendar year of initial onset of psychotic symptoms	-0.007	0.993 (0.976-1.012)	0.474
Ever hospitalized in psychiatric department (1=yes 2=no)	**-0.799**	**0.450 (0.310-0.653)**	**<0.001*****
Followed up by primary care providers in the recent year (1=yes 2=no)	0.197	1.218 (0.835-1.775)	0.306
Receiving free antipsychotic medication from NISPP system (1=yes, 2=no)	**-0.311**	**0.733 (0.518-1.037)**	**0.079***
Receiving free antipsychotic medication from DPF system (1=yes, 2=no)	-0.015	0.985 (0.664-1.462)	0.940
Community-level characteristics			
percent rural residents	1.347	3.845 (0.744-19.867)	0.108
percent ethnic minority residents	0.166	1.180 (0.606-2.298)	0.626
distance to the nearest psychiatric hospital (km)	-0.003	0.998 (0.991-1.004)	0.441
average GDP per capita of the county at time of survey	-0.046	0.955 (0.899-1.014)	0.133
community is part of a county officially classified as 'mountainous'	0.027	1.028 (0.650-1.625)	0.907
has psychiatric hospital within the county (1=yes 2=no)	-0.332	0.717 (0.458-1.123)	0.146
Interaction terms			
age*percent rural residents in community	**-0.027**	**0.974 (0.945-1.004)**	**0.086***

[a] 226 cases were removed from the analysis because they have missing values in one or more covariates

Rho (residual intraclass correlation) <0.001; Sigma_u =0.002 (95% CI: 2.72e-10-11322.61)

* p<0.1, ** p<0.01, *** p<0.001

2.5.5 collinearity

Most covariates in these models, e.g., age, gender, years of schooling, ethnicity, etc., are conceptually independent and exogenous. However, it is concerned that there might be relationship between demographic factors and three covariates, ever been

hospitalized in psychiatric department, receiving free medication from NISPP system, and receiving free medication from DPF system. We firstly check the relationship between the probability of ever being hospitalized in psychiatric hospital and other covariates in equation one using multilevel modelling. As shown in the table 2.14, only gender is a significant predictor of ever being hospitalized, and gender itself is a significant in equation four. Therefore, the collinearity between ever being hospitalized and other covariates is not a big issue.

Table 2.14. The relationship between the probability of ever being hospitalized and other covariates in equation two

	coefficient	p
Patient-level		
age in years	-0.006	0.667
years of schooling	0.041	0.274
gender (1=female 2=male)	0.266	0.043*
marital status (1=currently married, 2=other)	0.197	0.166
ethnicity (1=Han 2=ethnic minority)	0.071	0.717
Economic status of the family (1=impoverished, 2=not impoverished)	-0.009	0.967
Village-/neighborhood-level		
percentage of rural residents	0.427	0.653
distance to the nearest psychiatric hospital (km)	0.008	0.541
quadratic term of distance to the nearest psychiatric hospital (km)	<0.001	0.945
percentage of ethnic minority residents	-0.302	0.719
mountainous	0.087	0.609
psychiatric hospital within the county (1=yes 2=no)	0.193	0.333
average GDP per capita of the county when survey was conducted	0.041	0.171
Interaction terms		
years of schooling*percentage of rural residents	0.021	0.621
years of schooling*percentage of ethnic minority residents	-0.035	0.388
years of schooling*distance to the nearest psychiatric service	-0.0003	0.475
age*percentage of rural residents	-0.019	0.218
age*percentage of ethnic minority residents	0.003	0.836
age*distance to the nearest psychiatric service	-0.0002	0.225

[a] 240 cases are removed from the analysis because they have missing values in one or more covariates
Rho (residual intraclass correlation) = 0.038
Sigma_u =0.364 (95% CI: 0.147-0.902)
* p<0.1, ** p<0.01, *** p<0.001

The relationship between receiving free medication and other covariates in equation 4 is shown in table 2.15. Four variables have significant association with receiving free anti-psychotic medication from NISPP system: followed up by primary care doctors,

having psychiatric hospital within the county, distance to the nearest psychiatric hospital, and lower percentage of ethnic minority residents in the village/community, and two of these variables (having psychiatric hospital within the county and followed up by primary care doctors) are still significant in equation four. Therefore, the p-value of two coefficients, distance to the nearest psychiatric hospital and percentage of ethnic minority residents could be overestimated by equation 4. Followed up by primary care doctors is also significantly related to receive free medication from DPF system, and it is also significant in equation 4. Overall, there is be mild collinearity issue in equation 4 which is within an acceptable range.

Table 2.15. The relationship between the probability of receiving free anti-psychotic medication and other covariates in equation three

	Receiving free medication from NISPP system		Receiving free medication from DPF system	
	coefficient	p	coefficient	p
Patient-level				
age in years	-0.007	0.655	-0.017	0.138
years of schooling	0.009	0.713	0.015	0.380
gender (1=female 2=male)	0.093	0.646	0.188	0.168
marital status (1=currently married, 2=other)	0.132	0.551	0.042	0.778
ethnicity (1=Han 2=ethnic minority)	-0.245	0.509	0.024	0.911
Economic status of the family (1=impoverished, 2=not impoverished)	-0.541	0.132	-0.158	0.484
Followed up by primary care providers in the recent year (1=yes 2=no)	-0.676	0.005**	-0.269	0.077*
Village-/neighborhood-level				
percentage of rural residents	0.106	0.916	-0.319	0.617
distance to the nearest psychiatric hospital (km)	-0.018	0.002**	0.001	0.652
percentage of ethnic minority residents	1.946	<0.001***	-0.293	0.299
mountainous	-0.548	0.121	-0.159	0.439
psychiatric hospital within the county (1=yes 2=no)	1.183	0.016*	-0.113	0.597
average GDP per capita of the county when survey was conducted	0.065	0.231	-0.020	0.511
Interaction terms				
age*percentage of rural residents	0.003	0.850	0.008	0.509
Rho (residual intraclass correlation)	0.452		0.115	
Sigma_u	1.648 (1.252,2.168)		0.654 (0.400, 1.068)	
* p<0.1, ** p<0.01, *** p<0.001				

2.6 Discussion

Among the registered patients, most patients did initiate antipsychotic treatment (93.4%) and 74.0% of these patient had any use of antipsychotic medication in the last three months, which shows great improvement compared to the service utilization in 2001-2005 in China (only 60.4% of patients with psychotic disorder ever saw mental health professionals) (Phillips et al., 2009). The adherence rate with antipsychotic treatment in Ningxia (63.2%) is also comparable to that in developed counties (Lin et al., 2021; Kishi et al., 2016; Gutierrez-Casares et al., 2010). With the support of community mental health program, the treatment gap of severe mental illness has reduced dramatically especially in rural China.

The overall result of the determinants predicting mental health service use is summarized in the flowchart shown in Figure 2.20. One general summary conclusion is that individual demographic predictors are central to the initiation of antipsychotic medications, while community-level service-related factors are more closely related to sustaining regular use of antipsychotic medications.

Age plays an important role to the probability of initiating treatment and any use of medication in the three months prior to the survey, indicating older patients with psychosis not only have less chance to start treatment but also are more likely to interrupt their treatment. Years of schooling is significant predictor of the probability of ever seeking treatment in less urbanized area while it is significant in more urbanized areas and the areas with more ethnic minority population when predicting the probability of any use of medication in the last three months. Female patients have higher likelihood to initiate treatment and use medication in the three months prior to the survey, but males have better adherence among all the patients was recently using medication. Ethnic minority is significant factor of not ever seeking treatment, but it is not significant factor of recent medication use or adherence.

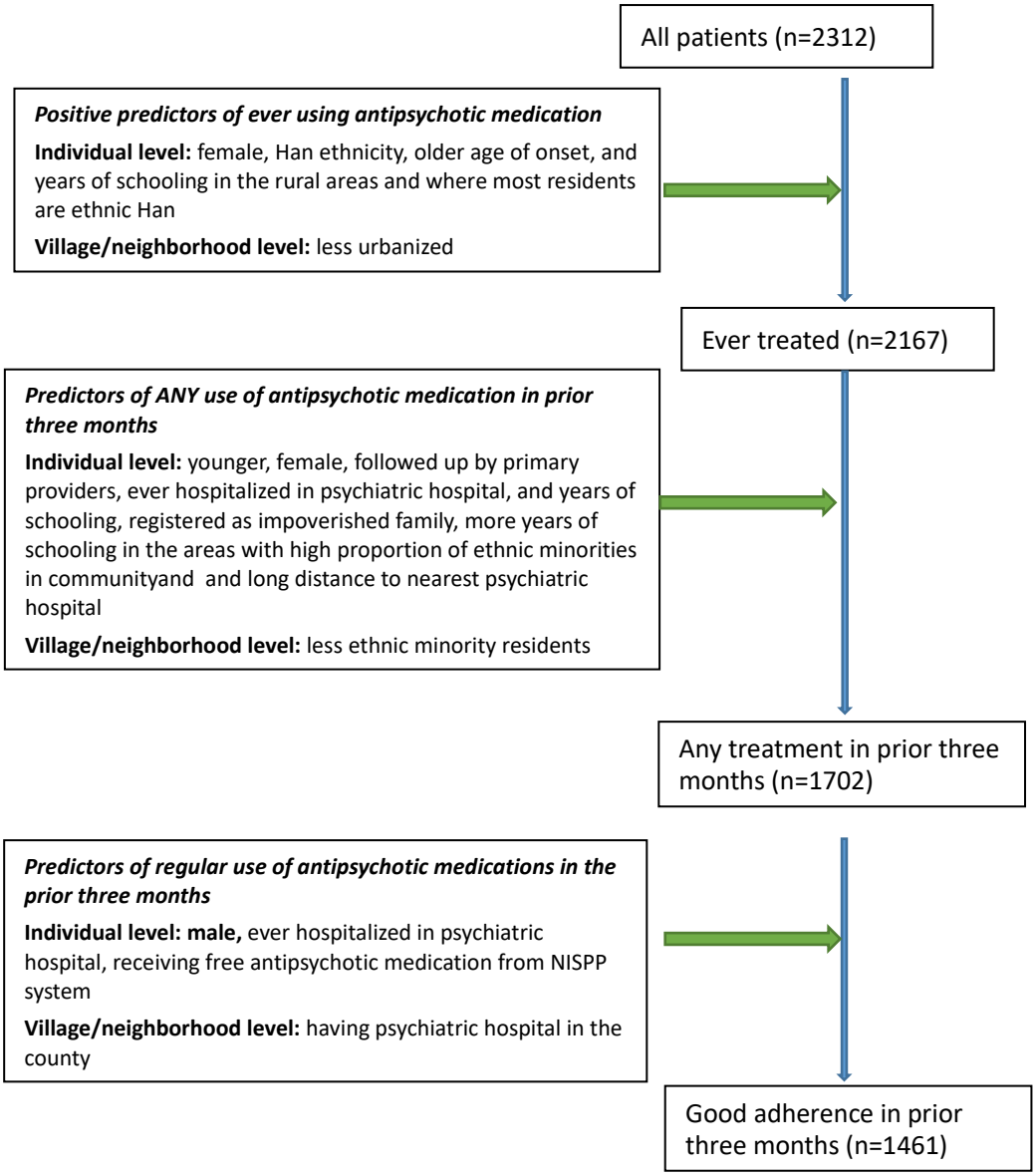

Figure 2.20 Summary of the significant predictors

This study also confirm that the characteristics of the location also have influence on the medication use. After controlling for other covariates, urban-rural disparity is extremely large of the probability to start treatment but the urban-rural difference of the probability of using medication in the last three months and having good adherence is not significant. The accessibility to psychiatric service is also related to the treatment

seeking behavior of patients: 1) distance to the psychiatric service is negatively related to the probability of initiating treatment where the distance to the nearest service is less than 90 km, but the distance is positively related to initiate treatment in the places where the distance to the nearest is greater than 90km; 2) in contrast, the distance is positively related to any use of medication in the last three months in the areas closer to the service but negatively related to the probability of medication use in the recent three months in the areas further from psychiatric service; 3) distance is not significantly related to adherence among all the patients using medication in the recent three months, but having psychiatric service within the county is significantly related to higher probability of having good adherence for the patients.

Prior service utilization has long-term effect on the adherence. It is shown in table 2.3 and 2.4 that prior hospitalization is the strongest predictor of the probability of using medication in the recent three months and having good adherence. The potential reason could be after a period of standard inpatient treatment, patients and their family members would have better understanding and knowledge about the disease and treatment and also become more inclined to maintain adherence to decrease the risk of relapse (Kikkert, & Dekker, 2017).

Though the government has implemented national community program to extent the outreach of mental health services, especially in the rural area, having psychiatric hospital within the county is still a significant factor of the probability of keeping adherence of medication use (OR=0.611, p=0.023) after adjusting for other covariates. The disparity of mental health service accessibility has not been completely eliminated by NISPP program.

Receiving free anti-psychotic medication from NISPP system is a positively related to better adherence of drug use, which indicates providing free medication could be an effective strategy to improve the adherence of patients with psychosis. However, it could also due to potential selection bias, i.e., the system selects the patients who are more inclined to maintain better adherence to distribute the anti-psychotics. The following up service provided by primary care doctors is a significant factor of any use

of medication in the three months prior to the survey but not related to medication adherence, which also could be explained from two aspects. It could due to following up by primary care doctors increases the use anti-psychotics by promoting knowledge of mental illness, but it could also because of the fact that the primary care doctors are more likely to only follow up the patients using medication.

Furthermore, it could be told from this study that disparity of mental health service use still exists under the implement of national mental health service. Younger patients and patients with better educational attainment are still more likely to initiate treatment and keep better adherence. The accessibility of mental health service is also related to medication use among these registered patients. For example, having psychiatric hospital within the county associated with higher probability of keep adherence after controlling for other covariates in the model, indicating that community mental health program has not achieved its ultimate goal, universal mental health service for all the patients living in the community regardless their accessibility to the psychiatric service from the hospital. Moreover, our study shows that patients living closer to the psychiatric hospital and in the county where there exists psychiatric hospital are more likely to receive free medication (appendix table 2.15) because there is not enough manpower to implement the community mental health (including delivering free medications, enrolling new patients to the system, etc.) in the places without psychiatric hospital. In this case, the disparity of mental health service use could be aggravated if we cannot increase the service providers in the areas lack of psychiatric hospital.

This study has two main limitations. First and foremost, this study conducted a relatively short interview, so the models had omitted variables issue, such as the income level of patients, mental health literacy of both patients and their guardians, stigma, and discrimination. Usually, income, mental health literacy, and knowledge have positive relationship with years of schooling, so the effect of years of schooling could be overstated due to these omitted variables. More importantly, this study is limited by its cross-sectional nature. We identify the current significant predictors of medication use

and disparity of service use between areas with and without psychiatric hospital, but we cannot make any causal inference or make the judgement whether national community mental health has improved the accessibility of mental health service and reduced the disparity of mental health service between areas with and without psychiatric hospital. We still need longitudinal data or construct longitudinal model to measure the effectiveness of the national-level community mental health program in China. Another limitation of study is selection bias of the study population. This study interviewed all the registered patients in NISPP system in these eight counties rather than all the patients with severe mental illness screened out by one epidemiological study, and the 'invisible patients' (who have disease but not yet registered by the system) could not be random missing. Therefore, we need to be careful when generalize the finding of this study to all the patients living in the community.

References:

Bland, R.C., Newman, S.C., & Orn, H. (1997). Help-seeking for psychiatric disorders. *Canadian Journal of Psychiatry*, 42, 935-42.

Charlson, F.J., Ferrari, A.J., Santomauro, D.F., Diminic, S., Stocking, E., & Scott JG. (2018). Global Epidemiology and Burden of Schizophrenia: Finding From the Global Burden of Disease Study 2016. *Schizophrenia Bulletin*, 44(6), 1195-1203.

Curto, M., Fazio, F., Ulivieri, M., Navari, S., Lionetto, L., & Baldessarini, R. (2021). Improving adherence to pharmacological treatment for schizophrenia: a systematic assessment. *Expert Opinion on Pharmacotherapy*. Published online 05 Feb 2021. https://doi.org/10.1080/14656566.2021.1882996

Eaton, W.W., Martins, S., Nestadt, G., Bienvenu, J., Clarke, D., & Alexandre, P. (2008). The burden of mental disorders. *Epidemiol Rev*, 30,1-14. doi: 10.1093/epirev/mxn011

Fenton WS, Blyler CR, Heinssen RK. (1997). Determinants of Medication Compliance in Schizophrenia: Empirical and Clinical Findings. *Schizophrenia Bulletin*, 23(4), 637-651

Gutierrez-Casares, J.R., Canas, F., Rodriguez-Morales, A., Hidalgo-Borrajo, R., & Alonso-Escolano, D. (2010). Adherence to treatment and therapeutic strategies in schizophrenic patients: the ADHERE study. *CNS Spectr,* 15(5), 327-37.

GBD 2016 Healthcare Access and Quality Collaborators. (2018). Measuring performance on the healthcare access and quality index for 195 countries and territories and selected subnational locations: a systematic analysis from the Global Burden of Disease Study 2016. *The Lancet,* **391**(10136):2236-2271. Doi: https://doi.org/10.1016/S0140-6736(20)30750-9.

He, H.R., Liu, Q.Q., Li, N., Guo, L.Y., Gao, F.J., Ling, B., … & Lyu, J. (2017) Trends in the incidence and DALYs of schizophrenia at the global, regional and national levels: results from the Global Burden of Disease Study 2017. *Epidemiology and Psychiatric Sciences*, 29(e91), 1–11. Doi: https://doi.org/10.1017/S2045796019000891

Kikkert, M.J., & Dekker, J. (2017). Medication Adherence Decisions in Patients with Schizophrenia. *Primary Care Companion CNS Disorder*, 19(6), 17n02182. Doi: https://doi.org/10.4088/PCC.17n02182

Kishi, T., Oya, K., & Iwata, N. (2016) Long-acting injectable antipsychotics for the prevention of relapse in patients with recent onset psychotic disorders: a systematic review and meta-analysis of randomized controlled trials. *Psychiatry Research*. 246, 750-755

Kohn, R., Saxena, S., Levav, I., & Saraceno, B. (2004) The treatment gap in mental health care. *Bulletin of the World Health Organization*, 82(11), 858-866.

Lin, D., Pilon, D., Zhdanava, M., Joshi, K., Lafeuille, M.H., Cote-sergent, A., … & Lefebvre, P. (2021). Medication adherence, healthcare resource utilization, and costs among Medicaid beneficiaries with schizophrenia treated with once-monthly paliperidone palmitate or once-every-three-months paliperidone palmitate. *Current Medical Research and Opinion*, 37(4), 675-683.

Lora, A., Kohn, R., Levav, I., McBain, R., Morris, J., & Saxena, S. (2012) Service availability and utilization and treatment gap for schizophrenic disorders: a survey in 50 low- and middle-income countries. *Bull World Health Organ*, 90, 47-54B. doi:10.2471/BLT.11.089284

Moreno-Kustner, B., Martin, C., & Pastpr, L. (2018). Prevalence of psychotic disorders and its association with methodological issues. A systematic review and meta-analyses. *Plos One*, 13(4), e0195687.

National Health Commission of the People's Republic of China. (2017). China Health Statistics Yearbook 2017. Peking Union Medical College Press.

Phillips, M.R., Zhang, J.X., Shi, Q.C., Song, Z.Q., Ding, Z.J., Pang, S.T., … & Wang Z.Q. (2009). Prevalence, treatment, and associated disability of mental disorders in four provinces in China during 2001-05: an epidemiological survey. *The Lancet,* 373(9680), 2041-2053.

Rabe-Hesketh, S., & Skrondal, A. (2012). Multilevel and Longitudinal Modeling Using Stata. Stata Press.

Shi, C.H., Ma, N., Wang, L.Y., Yi, L.L., Wang, X., Zhang WF, et al. (2019). Study of the Mental Health Resources in China. *Chinese Journal of Health Policy,* 12(2), 51-57.

Subramaniam, M., Abdin, E., Vaingankar, J., Sambasivam, R., Zhang, Y.J., Shafie, S., … & Chong, S.A. (2021) Lifetime prevalence and correlates of schizophrenia and other psychotic disorders in Singapore. *Front Psychiatry*. Published online 11 Mar 2021

Thornicroft, G., Szmukler, G., Mueser, K.T., & Krake, R.E. (2011). Oxford Textbook of Community Mental Health. Oxford University Press. New York

Thornicroft, G., Deb, T., & Henderson, C. (2016) Community mental health care worldwide: current status and further developments. *World Psychiatry*. 15(3), 276-286.

WHO. (2013). Mental Health Action Plan 2013-2020; World Health Organization: Geneva, Switzerland.

WHO. (2008). Mental Health Action Programme: scaling up care for mental, neurological, and substance use disorders.

Zhao, M.M., Ma, N., Wang, X., Wu, X.M., Zhang, W.F., Chen, R.Z., … & Lin, L. (2020). Community-Based Management and Treatment Services for Psychosis-China, 2019. *China CDC Weekly*, 41(2), 79

Chapter Three: Assessing social cognition in patients with schizophrenia and healthy controls using the Reading the Mind in the Eyes Test (RMET): a systematic review and meta-regression

Abstract

Background. Reading the Mind in the Eyes Test (RMET) is a widely used measure that assesses the theory of mind, one of four widely recognized components of social cognition. Most studies comparing RMET scores between patients with schizophrenia and healthy controls find lower scores in patients, but there is substantial heterogeneity across studies. Age and education may affect social cognition, but few studies about RMET assess the effect of these potential covariates – covariates that may have different effects in patients and healthy controls. We conducted a systematic review to clarify the state of knowledge about these issues.

Methods. Studies published before July 2020 that administered RMET to patients with schizophrenia or healthy controls were identified from three English-language databases (Pubmed, Web of Science, and Psycinfo) and two Chinese-language databases (China National Knowledge Infrastructure and Wanfang). The characteristics and results of identified studies were analyzed, and an 11-item quality assessment inventory was used to evaluate the methodological rigour of the studies.

Results. The electronic search identified 198 studies that provide data on 41 separate samples of patients with schizophrenia (total n=1836) and 197 separate samples of healthy controls (total n=23,675). Combining data from all studies, the pooled RMET score was 19.76 (95% CI 18.91-20.60) in patients with schizophrenia and 25.53 (95% CI 25.19-25.87) in healthy controls ($z=12.41$, $p<0.001$). Meta-analysis of 26 studies that directly compared RMET results in patients and healthy controls indicated that RMET scores in patients were 1.10 mean standard deviations lower than RMET scores in healthy controls ($z=12.32$, $p<0.001$). Meta-regression of 40 patient samples and 180 healthy control samples that include data on mean age and of 35 patient samples and

98 healthy control samples that include data on mean years of schooling found a negative association of age with RMET score and a positive association of years of schooling with RMET score, though these associations were only statistically significant in the healthy control samples. A secondary meta-analysis using a spline construction of the 180 healthy control samples that include data on mean age identified a non-monotonic relationship between age and RMET score – RMET scores increased with age before age 31 and decreased with age after age 31.

Conclusion. Patients with schizophrenia showed substantial deficits in theory of mind compared with healthy controls. The consistent relationship between increasing years of schooling and higher RMET scores in both patients and controls suggests that vocabulary (which is closely related to educational level) plays an important role in RMET performance. The non-monotonic relationship between age and RMET score in this large sample of healthy controls provides a potential new approach for assessing how theory of mind changes with aging.

Keywords: Reading the Mind in the Eyes Test, schizophrenia, healthy controls, years of education, aging, meta-regression

3.1 Introduction:

Individuals with schizophrenia generally exhibit neurocognitive deficits in multiple cognitive domains, including executive function, memory, attention, and problem-solving (Harvey, 2018; Mesholam-Gately, 2009; Sheffield, 2018). In addition to neurocognitive impairments, deficits in social cognition – the ability to learn social norms and perceive emotions and other social cues in interpersonal interactions – occur commonly in individuals with schizophrenia (Green 2019). The social cognition domain may be divided into four sub-domains: emotion processing, social perception, attributional style, and theory of mind (i.e., mentalizing) (Green, Horan, & Lee, 2015). Emotion processing and mentalizing are the two sub-domains that have been most studied in schizophrenia.

The Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) Consensus Cognitive Battery (MCCB) (Nuechterlein et al., 2008) is the most widely used battery to comprehensively assess cognition in schizophrenia. However, the test used to assess social cognition in this battery, the Mayer-Salovery-Caruso Emotional Intelligence Test (MSCEIT) (Mayer et al., 2002), expects respondents to interpret stories or vignettes about social situations that are unfamiliar to many respondents from non-Western cultures, particularly rural respondents, so it is frequently omitted in studies of cognition in schizophrenia (Stone et al., 2020, Deng et al., 2022).

The Reading the Mind in the Eyes Test (RMET) (Baron-Cohen, 2001) is a commonly used measure that evaluates an individual's theory of mind. Unlike other social-cognitive measures based on stories or vignettes (Mayer, 2003; Corcoran, 1995; Achim, 2012), RMET shows respondents pictures of the eye region of 36 faces with four adjectives describing different mental states and requires the participant to select the label that best describes the mental state of the individual portrayed in the picture. This 36-item English version of RMET has been translated into more than 20 languages and is widely used around the globe.

Most reported studies that use RMET to compare theory of mind in individuals with schizophrenia to that of healthy controls identify deficits in patients with schizophrenia. However, there are substantial differences in the magnitude of the reported deficits.

For example, Lopez-del-Hoyo and colleagues (2019) found that individuals with schizophrenia correctly categorized an average of 13 fewer pictures (out of 36) than healthy controls, while Okruszek and colleagues (2017) found that the patient group only misclassified an average of 2.4 more pictures than the control group. A meta-analysis combining the results of previous studies that compare RMET results between patients with schizophrenia and healthy controls is needed to provide a more precise estimate of the deficit in social cognition of individuals with schizophrenia.

Furthermore, very few of the available studies that assess social cognition of individuals with schizophrenia or healthy controls report multivariate analyses that explore the association between RMET results and important covariates, such as age and years of schooling – factors that could potentially explain the considerable heterogeneity of RMET performance among participants. Many of the studies were too small to support such an analysis, and some studies enrolled participants with a relatively narrow range of ages or similar levels of education, making it difficult to identify the effect of these covariates. One method of addressing this issue is to conduct a study-level meta-regression of all available studies.

Not only to quantify the impairment in patients with schizophrenia, this is the first known systematic review to identify all studies that use RMET to assess social cognition in separate samples of individuals with schizophrenia or healthy control subjects, not limiting ourselves to studies that include both of these groups. Meta-regression analyses of data from all of these studies will assess the relationship of age and level of education to RMET scores in healthy controls and individuals with schizophrenia. A meta-analysis of data from the subgroup of studies that directly compare RMET results in individuals with schizophrenia and healthy controls will estimate the magnitude of social cognitive deficits in patients with schizophrenia.

3.2 Method:

3.2.1 Search

The search algorithm identified some studies that include both patients with schizophrenia and healthy controls, other studies that include patients with

schizophrenia with no controls (or with other types of controls), and studies that include healthy controls compared to other types of patients (e.g., patients with autism, bipolar disorder, etc.).

We searched for relevant articles published before 15 July 2020 in three English-language databases (PubMed, Web of Science, and PsycINFO/EBSCO) and two Chinese-language databases (China National Knowledge Infrastructure [CNKI] and Wanfang). The search strategy of the title and abstract of documents included the following terms: ('RMET' or 'Reading the Mind in the Eyes' or 'Reading the Mind in the Eye') OR ('schizophrenia' AND 'eye test'). The detailed search strategy for each database is shown in the supplementary materials. Reference lists of the papers meeting eligibility criteria were individually searched to identify additional studies.

3.2.2 Eligibility criteria

Original research studies using the 36-item version of RMET that report the unadjusted RMET score (i.e., the number of correctly classified pictures) of patients with schizophrenia or healthy controls are included. There was no restriction on the method of diagnosing schizophrenia. Studies were excluded if the samples of individuals with schizophrenia or healthy controls included persons under 18 years of age or with a history of mental retardation, autism spectrum disorder, epilepsy, brain injury, brain disease, substance use disorder, or other mental disorder.

3.2.3 Selection of studies

Several reviewers (MAB, YRC, JT, XB, YC, JL, ZL, and QY) screened the titles and abstracts of studies identified in the electronic searches of the databases to decide whether they potentially met the eligibility criteria. Two independent reviewers had to agree on the classification of each article; disagreement was resolved by the senior author (FD). Full-text versions of the potentially eligible articles were then retrieved and independently reassessed by two reviewers (MAB, YRC, JT, XB, YC, JL, and ZL) to ensure that they met the inclusion criteria; disagreements about the final selection were resolved through discussion with the senior author (FD).

3.2.4 Data extraction

The following information about each selected article was entered in a pre-designed table:

- study characteristics (first author, title, journal, year of publication, and language of publication);
- type of study population(s) (patients with schizophrenia only, healthy subjects only, both patients with schizophrenia and healthy controls, or healthy controls compared to patients with other diagnoses);
- characteristics of the study population (country of test administration, source of participants, sampling method, inclusion or exclusion criteria of the study, diagnostic criteria employed to screen subjects, sample size);
- characteristics of included participants (gender, age, years of schooling, urban or rural residence, ethnicity, treatment status (of individuals with schizophrenia);
- language of RMET test;
- method of administering RMET (interviewer-completed, paper and pencil self-completion, computer-based self-completion, or online self-completion);
- RMET test results (mean and SD of RMET score and results of multivariate analyses if available), and
- (only from papers that include both patients with schizophrenia and healthy controls) crude and adjusted results of the comparison of RMET scores between patients with schizophrenia and healthy controls.

Two independent reviewers (MAB, YRC, JT, XB, YC, and QY) extracted data for each included study; the senior author (FD) made a final determination in cases where the two reviewers disagreed.

3.2.5 Quality assessment

The quality assessment scale developed for this study included the 11 items listed in Table 3.1. The list combined adapted versions of items used in the STROBE

(Strengthening the Reporting of Observational Studies in Epidemiology) statement (von Elm E, 2008) with items based on the authors' experience administrating the RMET test. Each item was coded as '1' or '0' based on whether the paper fulfilled the criteria specified in the item. Thus the range of the total quality score was 0 to 11. We categorized the overall quality score based on these scores as follows: 0-4='poor', 5-7='fair', 8-11='good'. Two reviewers independently assessed the quality of each paper (MAB, YRC, JT, XB, YC, JL, and ZL); disagreements in the scores of any of the eleven items were resolved by the senior author (FD).

Table 3.1 Operational definition of eleven items used to assess the quality of the included studies

No.	Item content	n (%)*
1	Introduction provides rationale and specific objectives (hypothesis) for study.	197 (99.5%)
2	Method clearly describes the study design.	193 (97.5%)
3	Method section describes study setting(s), location, and recruitment date.	22 (11.1%)
4	Diagnostic criteria, eligibility criteria, and recruitment process for individuals with schizophrenia and healthy controls are specified.	99 (50.0%)
5	Rationale for the sample size is provided.	10 (5.1%)
6	Describes all statistical methods used, and, if there is any missing data, how missing data is managed. (Assume no missing data if not mentioned.)	157 (79.3%)
7	Provides demographic characteristics of the sample that includes age and years of schooling.	126 (63.6%)
8	Reports numbers of individuals at each stage of study (e.g. numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing the assessment, and included in the analysis).	35 (17.7%)
9	Reports unadjusted mean of the number of correctly classified pictures (or % of correctly classified pictures) with standard deviation or confidence interval.	187 (94.4%)
10	Reports adjusted RMET score after controlling for age and years of schooling or reports the result of multivariate regression analyses using RMET score as the outcome variable that includes age and years of schooling as covariates.	3 (1.5%)
11	Discusses the limitations of the study.	163 (82.3%)

*number (percent) of the 198 studies included in the review that provide this information

3.2.6 Modelling strategy

The t-test was used to compare the study quality score between study samples of patients with schizophrenia and healthy controls and between samples using different language versions of the RMET (English vs non-English). The mean RMET score(s), the mean of the number of correctly classified pictures in each group of respondents, was used as the outcome variable (i.e., effect size) for each study.

The true overall effect size, $\hat{\theta}$, was computed as the weighted average of the study-specific effect sizes (equation one), where $\widehat{\theta_j}$ was the overserved size effect of study *j* (j=1, 2, …, K) and w_j was the weight of study *j*:

$$\hat{\theta} = \frac{\sum_{j=1}^{K} w_j \widehat{\theta_j}}{\sum_{j=1}^{k} w_j} \qquad (1)$$

In random-effect models, the observed effect size, $\widehat{\theta_j}$, deviated from the pooled effect of the population due to two error terms as below:

$$\widehat{\theta_j} = \theta_j + \varepsilon_j = \theta + u_j + \varepsilon_j \qquad (2)$$

The two components of error term, u_j and ε_j, were assumed to be independent with $\varepsilon_j \sim N(0, \hat{\sigma}_j^2)$ and $u_j \sim N(0, \tau^2)$, and the weight, w_j, equaled to $1/(\hat{\sigma}_j^2+\hat{\tau}^2)$.

DerSimonian-Laird random-effect models was used to calculate the estimator, τ^2, for estimating the weight in weight of each separately in patients with schizophrenia and healthy controls. The DerSimonian-Laird random-effect model is particularly useful when pooling samples that have heterogeneous results (DerSimonian R, 1986).The z-test was used to compare pooled estimates of RMET scores in patient and healthy control samples.

A random-effect model was used to compare the standard mean difference (SMD) of RMET scores between individuals with schizophrenia and healthy controls in the studies that included both types of respondents because the effect size estimates were heterogeneous. In this analysis, effect sizes for each group were weighted using the inverse variance method. Q statistics, which follow a chi-square distribution, were used to assess standardized within-study differences. The heterogeneity of estimates across studies was assessed using I^2, which represents the proportion of the variance in the

estimates due to heterogeneity (Higgins, 2003). A funnel plot was used to evaluate potential publication bias, and Egger's test was used to assess the small-size effect (Egger, 1997). Subgroup analysis evaluated the possible influence of the language of the administered RMET on the outcome.

Both univariate meta-regression and multivariate meta-regression assessed the association of age and years of schooling with the RMET score in individuals with schizophrenia and in healthy controls using equation 3, 4, and 5, where ε_j was the sampling error through which the effect size of a study deviated from its true effect and u_j denoted that even the true effect size of the study was only sampled from an overarching distribution of effect sizes.

$$\widehat{\theta_j} = \theta + \beta_{age} * age_j + u_j + \varepsilon_j \quad (3)$$

$$\widehat{\theta_j} = \theta + \beta_{edu} * edu_j + u_j + \varepsilon_j \quad (4)$$

$$\widehat{\theta_j} = \theta + \beta_{age} * age_j + \beta_{edu} * edu_j + u_j + \varepsilon_j \quad (5)$$

The meta-regression equations were estimated using two different methods: restricted maximum likelihood (Viechtbauer W, 2005) and bootstrap (Davison AC, 1997).

The mean age in the 180 samples of healthy controls that provided age data covered a wide range (from 18.7 to 71.7 years old), making it feasible to conduct a meta-regression with spline construction of age to identify a potential non-monotonic relationship between age and RMET score in both univariate and multivariate analyses.

$$\widehat{\theta_j} = \theta + \beta_{age} * age_j + \gamma * D_j | age > age_{knot} * (age - age_{knot}) + u_j + \varepsilon_j \quad (6)$$

Where $D_j = 1$ if age was greater than knot value.

For example, when age knot value was set as 30, the function was shown below:

$$\widehat{\theta_j} = \begin{cases} \theta + \beta_{age} * age_j + u_j + \varepsilon_j & if\ age \leq 30 \\ \theta + \beta_{age} * age_j + \gamma * (age - 30) + u_j + \varepsilon_j & if\ age > 30 \end{cases} \quad (7)$$

All ages from 25 to 45 were fitted as the knot value, and the model with lowest Akaike Information Criterion (AIC) was considered the best-fitted model.

Data were analyzed using STATA 17.0 version. Besides the figure showing the none-

linear regression between age and RMET score using linear spline models was generated by R, other figures were generated by STATA.

3.2.7 Registration

The protocol of this systematic review was registered on PROSPERO on 30 November 2020, before starting the title and abstract screening of the electronically identified studies (registration ID: CRD 42020216401).

3.3 Results

3.3.1 Selection of studies

As shown in the flowchart (Figure 3.1), the titles and abstracts of 1886 articles identified in English-language databases and 157 articles identified in Chinese-language bases were screened to identify potentially eligible papers. Based on this preliminary screening by two independent reviewers, the kappa values for potential inclusion were 0.72 for English articles and 0.77 for Chinese articles. The full text of potentially eligible articles (556 in English and 61 in Chinese) was then reviewed by two independent reviewers; the kappa value for inclusion based on this final screening was 0.61 for English articles and 0.52 for Chinese articles. After screening the electronically identified articles and identifying additional articles from the reference lists of selected articles, 198 studies were included in the analysis, 5 in Chinese and 193 in English. These 198 studies included 41 separate samples of patients with schizophrenia (with a total of 1836 patients) and 197 separate samples of healthy controls (with a total of 23,976 individuals). Only 26 (13.1%) of the studies (with 1455 patients with schizophrenia and 1087 healthy controls) directly compared RMET results in individuals with schizophrenia and healthy controls. Among the 41 samples of patients with schizophrenia, 8 (19.5%) used the English-language version of RMET, 3 (7.3%) used the Chinese-language version, 29 (70.7%) used other language versions, and 1 (2.4%) used two language versions (English and Korean). Among the 197 samples of healthy controls, 75 (38.1%) used the English-language version of RMET, 7 (4.1%) used the Chinese-language version, 110 (55.8%) used other language versions of RMET, 1 used

two language versions (English and Korean), and the language version used in 4 (2.0%) study samples was unknown. The detailed characteristics of these studies are shown in Table 3.2.

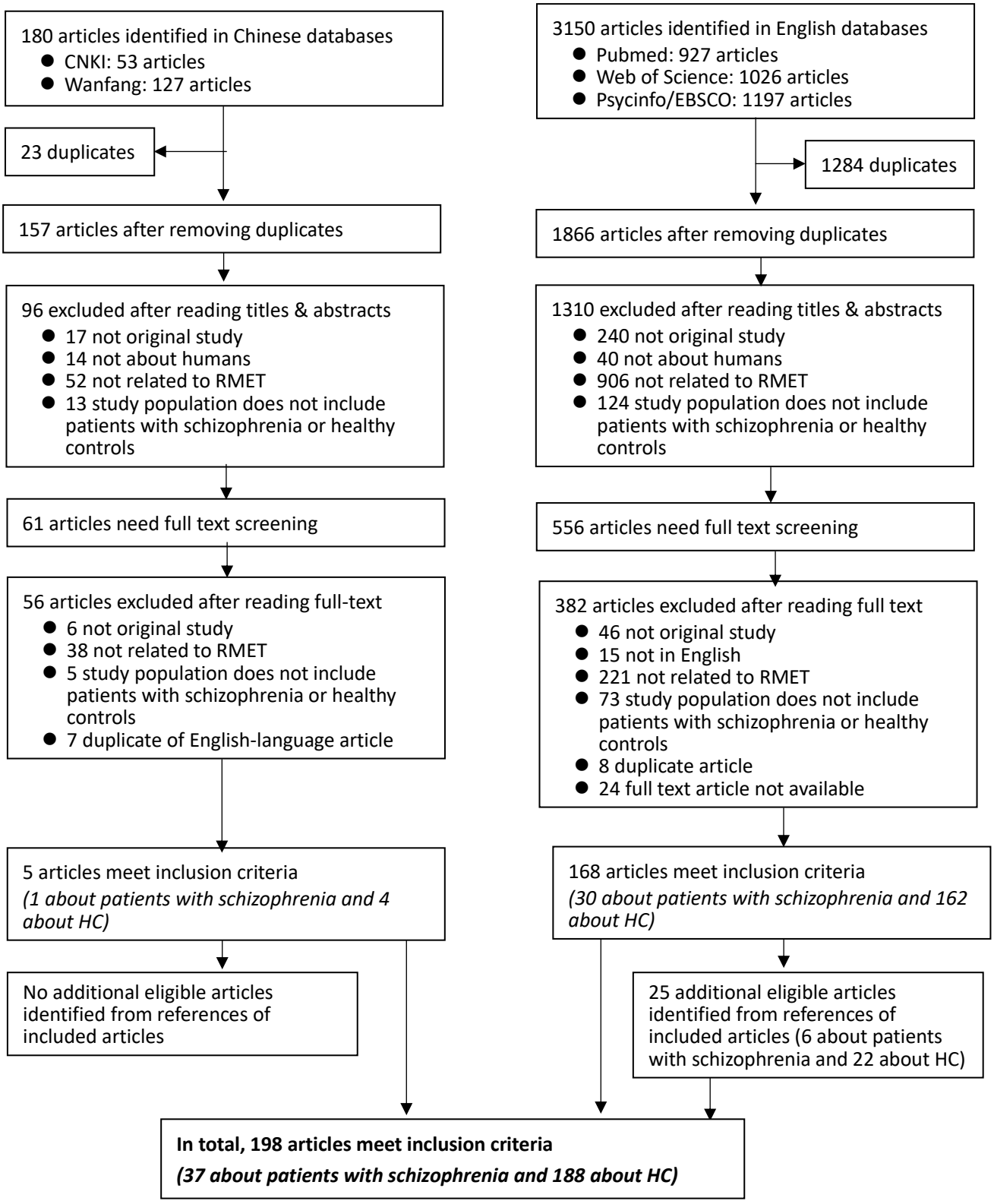

Figure 3.1 Flowchart of the identification of articles

Table 3.2 Characteristics of 40 samples of patients with schizophrenia and 197 samples of healthy controls reported in the 198 included studies

Author (year)	language	sample	Country	Language of RMET	Diagnostic criteria	Sample size	Mean (SD) RMET score	Quality score*
Baron-Cohen, S. (2001)	English	HC-A	US	English	NA	88	26.2 (3.6)	5
		HC-B	UK	English	NA	103	28.0 (3.5)	5
		HC-C	UK	English	NA	14	30.9 (3.0)	5
Craig, J.S. (2004)	English	HC	UK	English	NA	16	27.6 (4.3)	3
Kelemen, O. (2004)	English	HC	Hungary	Hungarian	DSM-IV	40	27.8 (5.0)	6
Havet-Thomassin, V. (2006)	English	HC	France	French	NA	17	27.6 (2.7)	5
Henry, J.D. (2006)	English	HC	UK	English	NA	17	25.9 (4.1)	4
Meyer, J. (2006)	English	HC	US	English	NA	142	28.9 (3.2)	5
Murphy, D. (2006)	English	SCH	UK	English	ICD-10	13	19.0 (SD NA)	3
Bora, E. (2007)	English	SCH	Turkey	Turkish	DSM V	58	18.2 (4.7)	8
Domes, G. (2007)	English	HC	Germany	German	NA	30	25.0 (2.9)	3
Kemmis, L. (2007)	English	HC	UK	English	NA	21	26.4 (5.1)	5
Murphy, D. (2007)	English	SCH	UK	English	NA	30	23.5 (6.1)	4
Sleesor, G. (2007)	English	HC	UK	English	NA	40	24.0 (4.1)	3
Hirao, K (2008)	English	SCH	Japan	Japanese	DSM-IV	20	19.2 (3.2)	7
		HC	Japan	Japanese	DSM-IV	20	24.3 (2.1)	7
Banati, M. (2009)	English	HC	Hungary	Hungarian	NA	35	25.2 (4.0)	4
Fertuck, E.A. (2009)	English	HC	US	English	DSM-IV	25	25.0 (3.6)	6
Harrison, A. (2009)	English	HC	UK	English	DSM-IV	20	30.6 (2.6)	5
Henry, J.D. (2009)	English	HC	Australia	English	na	30	28.7 (3.9)	5
Henry, J. D. (2009)	English	HC	Australia	English	NA	20	32.1 (1.9)	4
Pardini, M. (2009)	English	HC-A	US	Italian	NA	30	26.8 (2.8)	4
		HC-B	UK	Italian	NA	30	25.3 (3.2)	4
		HC-C	Italy	Italian	NA	30	23.5 (3.1)	4
		HC-D	Italy	Italian	NA	30	21.6 (2.2)	4
Rodrigues, S.M. (2009)	English	HC	US	English	NA	179	27.4 (3.5)	3
Russell, T.A. (2009)	English	HC	UK	English	NA	22	30.7 (1.9)	6
Schmidt, J.Z. (2009)	English	HC	Denmark	Bosnian and Danish	NA	56	26.0 (3.2)	7
Smeets, T. (2009)	English	HC	Netherlands and Germany	NA	NA	32	28.4 (1.1)	3
Couture, S.M. (2010)	English	SCH	US	English	DSM-IV	44	21.1 (5.7)	5
		HC	US	English	DSM-IV	41	25.0 (4.3)	5
de Achaval, D. (2010)	English	SCH	Argentina	Spanish	DSM IV	20	23.5 (5.4)	8
		HC	Argentina	Spanish	DSM IV	20	27.3 (3.8)	8
Geraci, A. (2010)	English	HC	Italy	Italian	NA	20	27.7 (4.1)	5
Gooding, D.C. (2010)	English	HC	US	English	NA	110	27.8 (3.0)	7
Kalbe, E. (2010)	English	HC	Germany	German	NA	28	25.6 (2.1)	5
Nietlisbach, G. (2010)	English	HC	Switzerland	German	NA	16	26.6 (2.8)	6
Lee, S.A. (2010)	English	HC	US	English	NA	96	27.6 (3.5)	4
Oldershaw, A. (2010)	English	HC	UK	English	DSM-IV	47	28.4 (3.5)	9
Schimansky, J. (2010)	English	SCH	Switzerland	German	ICD-10	40	23.1 (3.2)	5
		HC	Switzerland	German	NA	39	25.5 (2.6)	5
Tso, I.F. (2010)	English	HC	US	English	NA	33	26.4 (4.2)	7
Valla, J.M.(2010)	English	HC	US	English	NA	144	27.2 (3.3)	3

Table 3.2 Characteristics of 40 samples of patients with schizophrenia and 197 samples of healthy controls reported in the 198 included studies

Author (year)	language	sample	Country	Language of RMET	Diagnostic criteria	Sample size	Mean (SD) RMET score	Quality score*
Zhang, W.(2010)	Chinese	HC	China	Chinese	NA	126	20.4 (4.0)	4
Ahmed, F.S (2011)	English	HC	US	English	DSM-IV	123	27.3 (3.8)	5
Bai, Q.Y. (2011)	Chinese	HC	China	Chinese	NA	118	22.1 (3.9)	4
Cavallo, M.(2011)	English	HC	UK	English	na	21	27.0 (4.5)	5
Gooding, D.C. (2011)	English	HC	US	English	NA	68	27.2 (3.1)	6
Kim, Y.T. (2011)	English	HC	Korea	korean	DSM-IV	27	25.6 (3.4)	7
Kornriech, C. (2011)	English	HC	Belgium	NA	DSM-IV-TR	25	23 (2)	6
Philipp, A. (2011)	English	HC	France	French	NA	18	26.3 (2.2)	6
Petroni, A. (2011)	English	HC	Argentina	Spanish	NA	16	26.7 (0.9)	4
Stanford, A.D. (2011)	English	HC	US	English	NA	14	27.9 (4.4)	7
Wolkenstein, L. (2011)	English	HC	Germany	German	DSM-IV	20	23.6 (4.0)	6
Adenzato, M. (2012)	English	HC	Italy	Italian	NA	32	27.0 (3.3)	6
Cusi, A.M. (2012)	English	HC	Canada	English	DSM IV	25	26.2 (3.1)	7
Donohoe, G. (2012)	English	SCH	Ireland	English	DSM-IV	487	22.9 (5.7)	6
		HC	Ireland	English	NA	163	26.2 (4.2)	6
Geraci, A.(2012)	English	SCH-A	Italy	Italian	DSM-IV	13	17.5 (5.0)	5
		SCH-B	Italy	Italian	DSM-IV	8	18.1 (2.1)	5
		SCH-C	Italy	Italian	DSM-IV	9	20.1 (4.4)	5
		HC	Italy	Italian	NA	20	27.7 (4.1)	5
Honekopp, J. (2012)	English	HC	Germany	NA	NA	1896	24.4 (3.8)	4
Hysek, C. (2012)	English	HC	Switzerland	German	NA	48	24.4 (1.1)	8
Kenyon, M. (2012)	English	HC	UK	English	DSM-IV	57	28.3 (3.2)	7
Kother, U. (2012)	English	HC	Germany	German	MINI	30	25.5 (3.8)	7
Medina-Pradas, C. (2012)	English	HC	Spain	Spanish	DSM-IV-TR	39	27.2 (2.7)	7
Nejati, V. (2012)	English	HC	Iran	Persian	DSM-IV	50	20.7 (2.8)	6
Pentaraki, A. D. (2012)	English	SCH	Greece	Greek	DSM-IV-TR	21	19.1 (4.1)	3
Quintana, D.S. (2012)	English	HC	Australia	English	NA	65	27.4 (3.1)	7
Schilling, L. (2012)	English	HC	Germany	German	MINI	27	25.7 (3.9)	5
Szanto, K. (2012)	English	HC	US	English	DSM-IV	28	24.5 (5.1)	7
Barrera, A. (2013)	English	HC	Argentina	Spanish	DSM-IV/ICD-10	12	23.8 (3.0)	6
Buhlmann, U. (2013)	English	HC	Germany	German	NA	31	25.7 (3.4)	6
Caletti, E. (2013)	English	SCH	Italy	Italian	DSM-IV-TR	30	19.6 (4.6)	6
		HC	Italy	Italian	NA	18	25.7 (3.7)	6
De Los Reyes, A. (2013)	English	HC	US	English	NA	50	25.2 (3.6)	5
De Sampaio, F.T.P (2013)	English	HC	Argentina	Spanish	DSM-IV	24	25.9 (2.9)	7
Dehning, S. (2013)	English	HC	Germany	English	NA	126	22 (4.3)	5
Dehning, S. (2013)	English	HC	Germany and Ethiopia	English	NA	257	16.9 (SD NA)	4
Fernández-Abascal, E.G. (2013)	English	HC	Spain	Spanish	NA	358	27.2 (3.6)	6
Giusti, L. (2013)	English	SCH	Italy	Italian	NA	20	17.6 (5.9)	7
		HC	Italy	Italian	MINI	17	26.2 (2.9)	7
Pardini, M. (2013)	English	HC	Italy	Italian	NA	4150	24.6 (2.1)	7
Poletti, M. (2013)	English	HC	Italy	Italian	NA	20	21.7 (3.0)	6
Poletti, M. (2013)	English	HC	Italy	Italian	NA	35	22.1 (4.2)	5

Table 3.2 Characteristics of 40 samples of patients with schizophrenia and 197 samples of healthy controls reported in the 198 included studies

Author (year)	language	sample	Country	Language of RMET	Diagnostic criteria	Sample size	Mean (SD) RMET score	Quality score*
Purcell, A.L. (2013)	English	HC	US	English	DSM-IV-TR	28	27.7 (SD NA)	5
Preller, K.H. (2013)	English	HC	Switzerland	German	DSM-IV	68	25.5 (3.8)	7
Thaler, N.S. (2013)	English	SCH	US	English	DSM-IV	30	18.7 (6.1)	8
		HC	US	English	DSM-IV	24	28.4 (2.8)	8
Thoma, P. (2013)	English	HC	Germany	German	ICD-10	20	24.5 (4.0)	8
Whitton, A. E. (2013)	English	SCH	Australia	English	DSM-IV	34	24.0 (5.7)	3
		HC	Australia	English	NA	44	26.7 (4.7)	3
Balogh, N. (2014)	English	SCH	Hungary	Hungarian	DSM-IV	43	17.9 (5.2)	6
		HC	Hungary	Hungarian	NA	41	24.1 (3.8)	
Bedwell, J.S. (2014)	English	HC	United States	English	DSM-IV-TR	686	24.0 (4.3)	7
Csukly, G. (2014)	English	HC	Hungary	Hungarian	DSM IV	29	25.3 (3.7)	6
		SCH-A	Hungary	Hungarian	DSM-IV	30	19.1 (4.9)	6
		SCH-B	Hungary	Hungarian	DSM-IV	28	21.3 (5.2)	6
Dal Monte, O. (2014)	English	HC	US	English	NA	29	25.2 (0.6)	7
Demartini, B. (2014)	English	HC	UK	English	DSM IV	34	24.2 (3.9)	6
Goddard, E. (2014)	English	HC	UK	English	NA	42	26.8 (4.0)	6
Gong, P.Y. (2014)	English	HC	China	Chinese	NA	322	21.2 (4.0)	5
Lam, B. Y (2014)	English	SCH	Hong Kong, China	Chinese	DSM-IV,	58	19.1 (4.1)	8
		HC	Hong Kong, China	Chinese	NA	61	22.7 (3.1)	8
Montag, C. (2014)	English	SCH-A	Germany	German	DSM-IV	16	20.6 (5.5)	9
		SCH-B	Germany	German	DSM-IV	19	20.7 (5.9)	9
Nandrino, J. (2014)	English	HC	France	French	NA	30	24.0 (3.3)	6
Prevost, M. (2014)	English	HC-A	Canada	French	NA	127	24.8 (3.8)	6
		HC-B	Canada	French	NA	139	25.6 (5.4)	6
Reynolds, M.T. (2014)	English	HC	Australia	English	DSM-IV-TR	20	28.8 (2.7)	5
Woolley, J. D., (2014)	English	HC	US	English	DSM-IV	31	26.2 (3.6)	7
Youssef, F.F. (2014)	English	HC	West Indies	NA	NA	655	25.7 (3.9)	4
Ainley, V. (2015)	English	HC	UK	English	NA	97	26.6 (6.1)	4
Baron-Cohen, S. (2015)	English	HC-A	US	English	DSM-IV or ICD-10	152	25.5 (4.6)	7
		HC-B	UK	English	DSM-IV or ICD-10	168	27.4 (3.4)	7
Boucher, O. (2015)	English	HC	Canada	French	NA	20	26.2 (3.6)	7
Cabinio, M. (2015)	English	HC	Italy	Italian	NA	36	26 (3.9)	4
Carré, J. M. (2015)	English	HC	Canada	English	NA	30	26.0 (3.5)	6
Eddy, C.M. (2015)	English	HC	UK	English	NA	26	26.7 (3.1)	6
Enrici, I. (2015)	English	HC	Italy	Italian	NA	25	23.4 (4.7)	5
Fisher, A.D. (2015)	English	HC	Italy	Italian	DSM-5	43	25.6 (4.5)	7
Jelsone-Swain, L. (2015)	English	HC	US	NA	NA	17	25.4 (2.3)	8
Launay, J. (2015)	English	HC-A	US	English	NA	279	26.9 (4.8)	6
		HC-B	UK	English	NA	26	25.9 (6.2)	6
Lawson, R. (2015)	English	HC	UK	English	NA	152	26 (3.8)	3
Ma, J. (2015)	Chinese	HC	China	Chinese	NA	95	23.7 (4.7)	6
Melchers, M. (2015)	English	HC	Germany	German	NA	108	24.6 (2.5)	6
Moieni, M., (2015)	English	HC	US	English	DSM-IV	109	28.2 (0.5)	9

Table 3.2 Characteristics of 40 samples of patients with schizophrenia and 197 samples of healthy controls reported in the 198 included studies

Author (year)	language	sample	Country	Language of RMET	Diagnostic criteria	Sample size	Mean (SD) RMET score	Quality score*
Radke, S. (2015)	English	HC	Netherlands	Dutch	NA	24	24.4 (4.5)	7
Robinson, L. J. (2015)	English	HC	UK	English	NA	28	26.8 (3.5)	3
Scherzer, P. (2015)	English	SCH	Canada	French	DSM-IV	21	20.2 (2.9)	7
		HC	Canada	French	NA	29	21.7 (4.0)	7
Schuwerk, T. .(2015)	English	HC	Germany	German	NA	17	25.5 (3.2)	5
Tella, M.D (2015)	English	HC	Italy	Italian	NA	41	26.8 (3.6)	6
Weisman, O. (2015)	English	HC	Singapore	English	NA	1463	25.9 (4.3)	5
Zilioli, S. (2015)	English	HC	US	English	NA	469	26.8 (3.5)	6
Ayesa-Arriola, R. (2016)	English	HC	Spain	Spanish	NA	159	23.8 (4.3)	8
Besnard, J. (2016)	English	HC	France	French	NA	30	21.3 (3.9)	6
Brambilla, M. (2016)	English	SCH	Italy	Italian	DSM IV	32	20.8 (4.8)	9
Bora, E. (2016)	English	SCH	Turkey	Turkish	DSM-IV	54	14.2 (3.9)	6
		HC	Turkey	Turkish	NA	27	19.4 (2.6)	6
Burke, T. (2016)	English	HC	Ireland	English	NA	50	25.2 (4.9)	6
El Haj, M. (2016)	English	HC-A	US	French	NA	40	26.8 (4.5)	6
		HC-B	UK	French	NA	36	22.8 (4.5)	6
Filippo, S.G. (2016)	English	HC	Italy	Italian	NA	20	28 (5.0)	5
Haag, S. (2016)	English	HC	Germany	German	NA	29	23.6 (4.0)	8
Heitz, C. (2016)	English	HC	France	French	NA	16	23.9 (2.8)	7
Hoche, F. (2016)	English	HC	US	English	NA	57	27.9 (3.3)	5
Jankowiak-Siuda, K. (2016)	English	HC	Poland	Polish	NA	325	25.0 (4.5)	5
Jermakow, N. (2016)	English	HC	Poland	Polish	NA	60	24.8 (4.3)	6
Melchers, M. (2016)	English	HC	Germany	German	NA	716	23.0 (3.3)	5
Okruszek, L. (2016)	English	HC	Poland	Polish	NA	47	25.7 (4.1)	7
Petersen, R. (2016)	English	HC	Australia	English	DSM-IV	20	29.1 (3.2)	8
Pino, M.C. (2016)	English	HC	Italy	Italian	DSM-IV	23	24.4 (6.3)	6
Rominger, C. (2016)	English	HC	Austria	German	NA	20	23.5 (3.9)	4
Sabater, A. (2016)	English	HC	Spain	Spanish	NA	25	25.6 (2.4)	5
Schneider, A. (2016)	English	HC	US	English	NA	65	28.0 (4.1)	6
Wang, X. (2016)	English	HC	China	Chinese	NA	268	23.6 (3.5)	6
Bodnar, A. (2017)	English	HC	Poland	Polish	NA	25	28.1 (3.4)	6
Charernboon, T. (2017)	English	SCH	Thailand	Thai	DSM-5	36	18.9 (4.4)	10
		HC	Thailand	Thai	NA	36	23.5 (4.4)	10
Chen, K. W. (2017)	English	SCH	Taiwan	Traditional Chinese	DSM V	53	19.5 (3.4)	7
DeAngelo, G. (2017)	English	HC	US	English	NA	141	21.7 (3.7)	2
Dulau, C. (2017)	English	HC	France	French	NA	65	24.0 (3.2)	7
Enrici, I. (2017)	English	HC	Italy	Italian	NA	20	24.2 (4.7)	5
Erdeniz, B. (2017)	English	SCH	Turkey	Turkish	DSM-IV-TR	23	16.0 (5.6)	6
		HC	Turkey	Turkish	NA	23	24.5 (3.9)	6
Gavilán, J.M. (2017)	English	HC	Spain	English	NA	95	25.1 (3.3)	5
Helle, S. (2017)	English	SCH	US	English	DSM-IV	87	21.1 (5.4)	7
Hotier, S. (2017)	English	HC	France	French	NA	36	24.4 (SD NA)	5
Leppanen, J. (2017)	English	HC	UK	English	DSM-5	29	25.8 (3.5)	8
Martinez-Sanchez, F. (2017)	English	HC	Spain	Spanish	NA	1398	26.8 (3.4)	6

Table 3.2 Characteristics of 40 samples of patients with schizophrenia and 197 samples of healthy controls reported in the 198 included studies

Author (year)	language	sample	Country	Language of RMET	Diagnostic criteria	Sample size	Mean (SD) RMET score	Quality score*
Newbury-Helps, J. (2017)	English	HC	UK	English	NA	42	25.7 (0.8)	5
Preti., A. (2017)	English	HC	Italy	Italian	NA	200	24.8 (4.2)	6
Raimo, S. (2017)	English	HC	Italy	Italian	DSM-V	40	29.0 (0.6)	7
Sanvicente-Vieira, B (2017)	English	HC	Brazil	Portuguese	NA	30	25.1 (2.7)	7
Sun, F. (2017)	Chinese	HC	China	Chinese	NA	49	23.2 (4.1)	5
Tylec, A., (2017)	English	HC	Poland	Polish	NA	50	25.7 (5.2)	7
Zabihzadeh, A. (2017)	English	HC	Iran	Persian	DSM-IV	25	24.9 (4.8)	7
Anupama, V. (2018)	English	HC	India	Kannada	NA	20	28.9 (2.0)	5
Atoui, M. (2018)	English	SCH	Lebanon	Lebanese	DSM-5	22	22.4 (5.8)	7
		HC	Lebanon	Lebanese	NA	21	26.6 (2.6)	7
Aydin, O. (2018)	English	SCH	Turkey	Turkish	DSM-IV-TR	34	16.5 (4.5)	6
		HC	Turkey	Turkish	NA	31	20.7 (4.8)	6
Berenson, K.R. (2018)	English	HC	US	English	DSM-IV	60	27.7 (0.5)	6
Coyle, T.R. (2018)	English	HC	US	English	NA	249	27.6 (3.5)	5
Eddy, C.M. (2018)	English	HC	UK	English	NA	28	27.4 (0.6)	6
Espinós, U. (2018)	English	HC	Spain	Spanish	DSM-IV-R	112	27.4 (3.2)	6
Fossati, A. (2018)	English	HC	Italty	Italian	NA	193	24.9 (3.9)	5
Grainger, S.A. (2018)	English	HC	Australia	English	NA	50	25.9 (1.4)	6
Khorashad, B. S. (2018)	English	HC	Iran	Persian	DSM-V	104	22.3 (2.7)	7
Lee, H.S. (2018)	English	SCH	US and Korea	Korean	DSM-IV	60	20.9 (5.4)	9
		HC	US and Korea	Korean	DSM-IV	60	26.7 (4.0)	9
Navarra-Ventura, G (2018)	English	HC	Spain	Spanish	NA	40	29.0 (2.4)	6
Lopez-Navarro, E. (2018)	English	SCH	Spain	Spanish	DSM-IV-TR	54	17.4 (4.5)	5
Pinkham, A.E. (2018)	English	HC	US	English	NA	154	24.8 (4.3)	7
Redondo, I. (2018)	English	HC	Spain	Spanish	NA	433	15.4 (2.2)	5
Uhlmann, A. (2018)	English	HC	South Africa	English	DSM-IV	21	24.1 (5.0)	7
Wang, Y.R. (2018)	Chinese	SCH-A	China	Chinese	DSM V	24	18.9 (4.8)	10
		SCH-B	China	Chinese	DSM V	22	17.2 (6.4)	10
Yang, C.C. (2018)	English	HC	UK	NA	NA	23	24.6 (4.4)	7
Adenzato, M. (2019)	English	HC	UK	English	NA	20	22.9 (3)	6
Altunbas, F.D. (2019)	English	HC	Turkey	Turkish	NA	30	31.0 (2.9)	5
Bayliss, L. (2019)	English	HC	Mexico	Spanish	NA	12	24 (2.7)	5
Black, J.E. (2019)	English	HC	US	English	NA	591	26.3 (5.3)	5
Budak, E.A. (2019)	English	HC	Turkey	Turkish	NA	60	25.6 (3.9)	7
Dalkner, N. (2019)	English	HC	Austria	German	NA	79	23.0 (4.2)	5
Duque-Alarcon, X. (2019)	English	HC	Mexico	Spanish	NA	15	25.7 (4.3)	6
Giordano, M. (2019)	English	HC	Mexico	Spanish	NA	116	25.6 (3.8)	6
Hartman, L.I. (2019)	English	SCH	Canada	English	DSM-IV-TR	44	21.2 (5.4)	7
		HC	Canada	English	DSM-IV-TR	62	25.2 (5.4)	7
Lopez-del-Hoyo, Y. (2019)	English	SCH	Spain	Spanish	DSM-IV or ICD-10	30	13.1 (6.7)	7
		HC	Spain	Spanish	NA	30	26.6 (4.7)	7

Table 3.2 Characteristics of 40 samples of patients with schizophrenia and 197 samples of healthy controls reported in the 198 included studies

Author (year)	language	sample	Country	Language of RMET	Diagnostic criteria	Sample size	Mean (SD) RMET score	Quality score*
Romosan, A.M. (2019)	English	HC	Romania	Romanian	NA	51	24.9 (4.6)	6
Sacchetti, S. (2019)	English	HC	UK	English	NA	51	28.3 (3.2)	8
Simon, M. (2019)	English	HC	Hungry	Hungarian	NA	32	26.7 (3.1)	7
Tapajóz, F. (2019)	English	HC	Argentina	Spanish	NA	17	26.5 (3.0)	6
Zegarra-Valdivia, J.A. (2019)	English	HC	Perú	Spanish	DSM-5	20	24.0 (3.3)	4
Charernboon, T. (2020)	English	SCH	Thailand	Thai	DSM V	50	19.1 (4.0)	8
		HC	Thailand	Thai	DSM V	50	23.4 (4.1)	8
Cotelli, M. (2020)	English	HC-A	Italy	English	NA	32	24.3 (0.5)	5
		HC-B	Italy	English	NA	30	22.1 (0.4)	5
Eddy, C.M. (2020)	English	HC	UK	English	NA	176	25.5 (1.5)	7
Ignatova, V.G. (2020)	English	HC	India	NA	NA	36	24.1 (3.9)	5
Keri, S. (2020)	English	SCH	Hungary	Hungarian	DSM-V	32	19.1 (4.3)	6
		HC	Hungary	Hungarian	NA	32	23.6 (4.9)	6
Kilic, F. (2020)	English	HC	Turkey	Turkish	DSM-IV-TR	35	23.2 (2.6)	7
McPhilemy, G. (2020)	English	HC	Ireland	English	DSM-V-TR	49	26.5 (3.8)	6
Pahnke, R. (2020)	English	HC	Germany	German	NA	119	25.6 (0.4)	5
Unal-Aydin, P. (2020)	English	HC	Sarajevo, Bosnia and Herzegovina	NA	NA	337	23.3 (3.3)	7
García-Fernández, L. (2020)	English	SCH-A	Spain	Spanish	DSM V	43	23.7 (4.4)	8
		HC-A	Spain	Spainish	MINI	50	26.7 (3.3)	8
		SCH-B	Spain	Spanish	DSM V	44	20.4 (4.8)	8
		HC-B	Spain	Spainish	MINI	52	24.8 (4.8)	8
Wilu Wilu, A. (2021)	English	HC	France	French	NA	28	27.5 (3.7)	5

HC, healthy control; SCH, patients with schizophrenia; NA, not available

*quality score assessed by study authors based on 11 items listed in Table 1 (total score ranges from 0 to 11)

3.3.2 Quality evaluation

Among the 41 samples of patients with schizophrenia included in the 198 papers, one reported a mean RMET score without an accompanying standard deviation (or standard error), and five did not include data on the mean educational level of participants. Among the 197 samples of healthy controls included in the 198 papers, four reported mean RMET scores without an accompanying standard deviation, 17 did not include data on the mean age of participants, and 99 did not include data on the mean educational level of participants.

The items used to assess study quality are shown in Table 3.1.

The results of the quality assessment of the 198 included studies are shown in the last column of Table 3.2. The total quality score (theoretical range 0-11) varied from 2 to 10. The mean (sd) quality score of all papers was 5.9 (1.4); 28 (14.1%) papers were classified as 'poor quality' (score=0-4), 148 (74.7%) as 'fair quality' (score=5-7), and 19 (9.6%) as 'good quality' (score=8-11). Among the 11 separate items, only five items were present in more than 75% of studies (items 1, 2, 6, 9, and 11 shown in Table 3.1). Four items were *absent* in more than 75% of the studies: description of study setting (item 3), rationale for sample size (item 5), number of study drop-outs (item 8), and adjustment of RMET results (item 10).

When assigning the quality assessed for the paper as a whole to each of the included samples in each paper, the overall mean quality score for the 238 samples was 6.0 (1.5); 22 (13.5%) poor quality, 173 (72.7%) fair quality, and 33 (13.9%) good quality. The mean quality score of the 41 samples of patients with schizophrenia was significantly higher than that of the 197 samples of healthy controls [6.7 (1.8) vs 5.9 (1.4); t=3.41, p<0.001]. The mean quality score in the 149 samples administered non-English versions of RMET was significantly higher than that of the 83 samples administered the English version of RMET [6.3 (1.4) vs 5.6 (1.5); t=3.13, p=0.002].

3.3.3 Pooled RMET scores of patients with schizophrenia and healthy controls

The pooled RMET scores in patients with schizophrenia and healthy controls are shown in Figures 3.2 and 3.3. Based on the results of 23,619 healthy controls reported in 193 separate study samples that provided both the mean and standard deviation of RMET scores, the pooled RMET score in healthy controls was 25.53 (95% CI, 25.19-25.86); this was significantly higher than the pooled RMET score of 19.76 (95% CI, 18.91-20.60) for 1823 patients with schizophrenia reported in 40 separate study samples that provided both the mean and standard deviation of RMET scores (z=12.41, p<0.001).

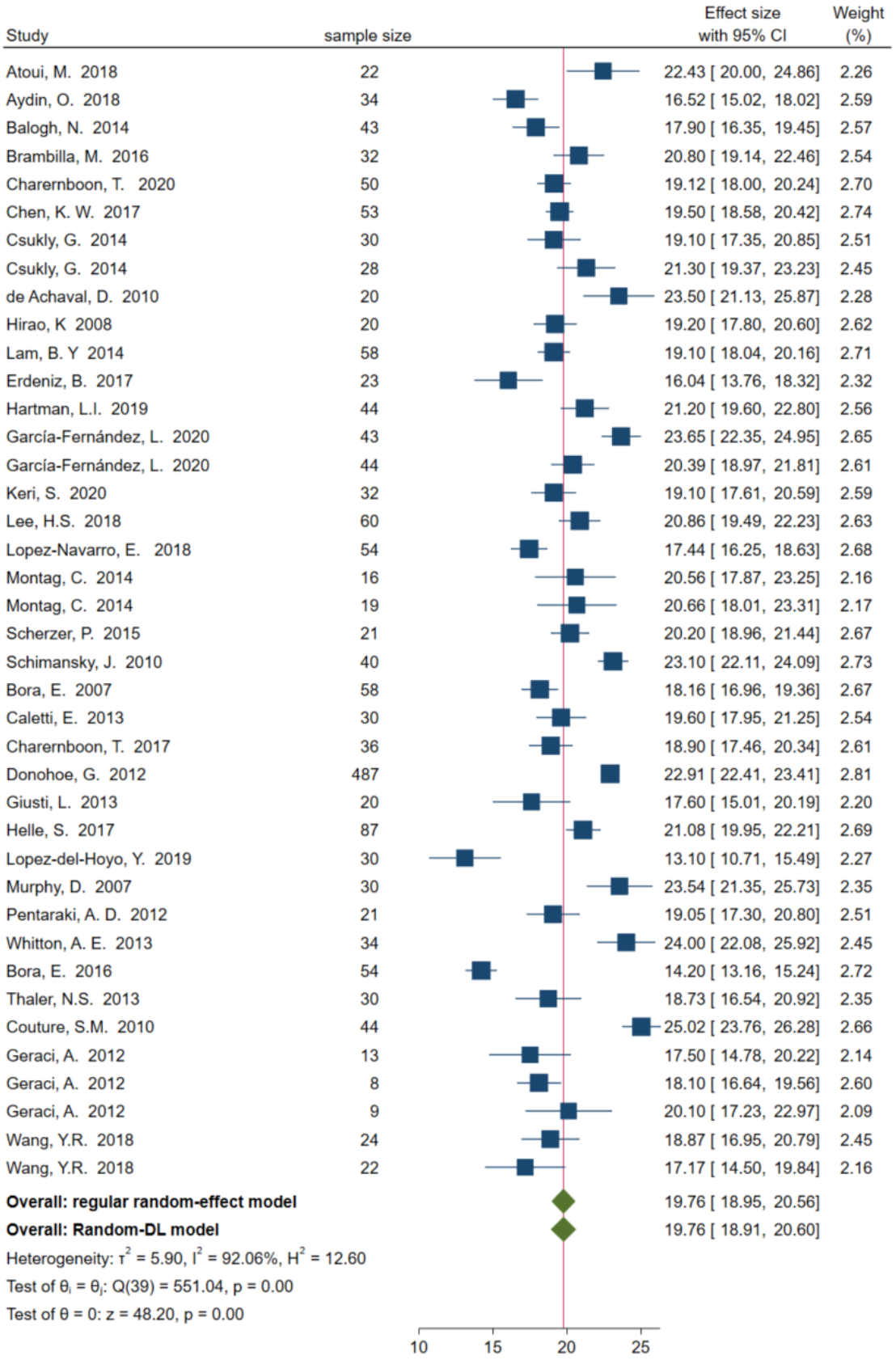

Figure 3.2 Pooled estimates of mean RMET scores in samples of patients with schizophrenia

(including 40 separate samples reported in 36 different papers with a total sample size of 1823 individuals with schizophrenia)

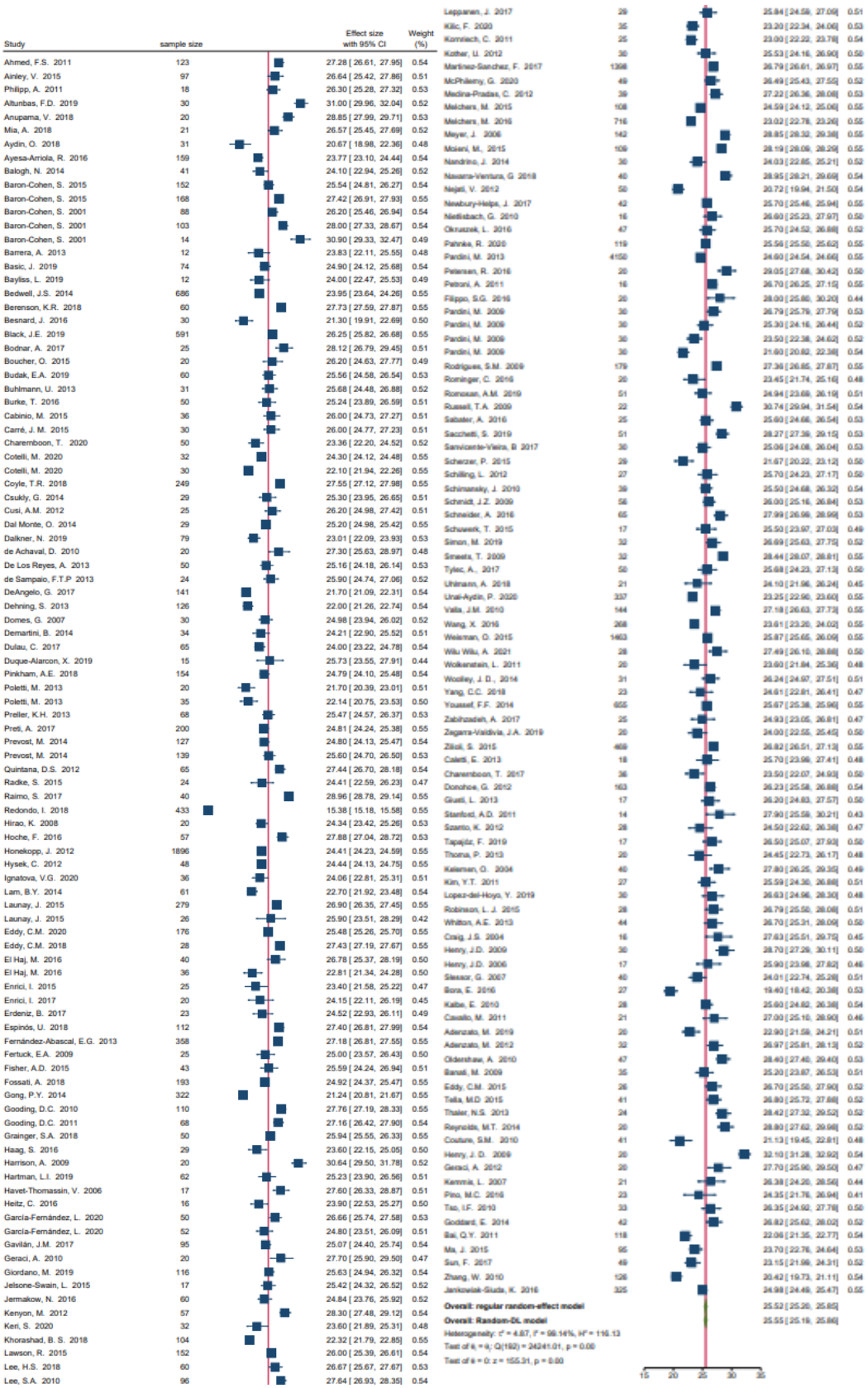

Figure 3.3 Pooled estimates of mean RMET scores in samples of healthy controls

(including 193 separate samples reported in 185 different papers with a total sample size of 23,619 individuals)

3.3.4 Direct comparison of RMET results between patients with schizophrenia and healthy controls

Among the 26 studies that directly compared mean RMET scores of patients with schizophrenia and healthy controls, only one study (Scherzer P, 2015) did not find a statistically significant difference between the two groups; all other studies reported significantly lower mean RMET scores in the patient group. As shown in Figure 3.4A, the pooled standard mean difference for the 26 studies estimated by a random-effect meta-analysis model indicated that the RMET scores in patients with schizophrenia were 1.10 standard deviations lower than the RMET scores in patients with schizophrenia (z=-12.32, p<0.001).

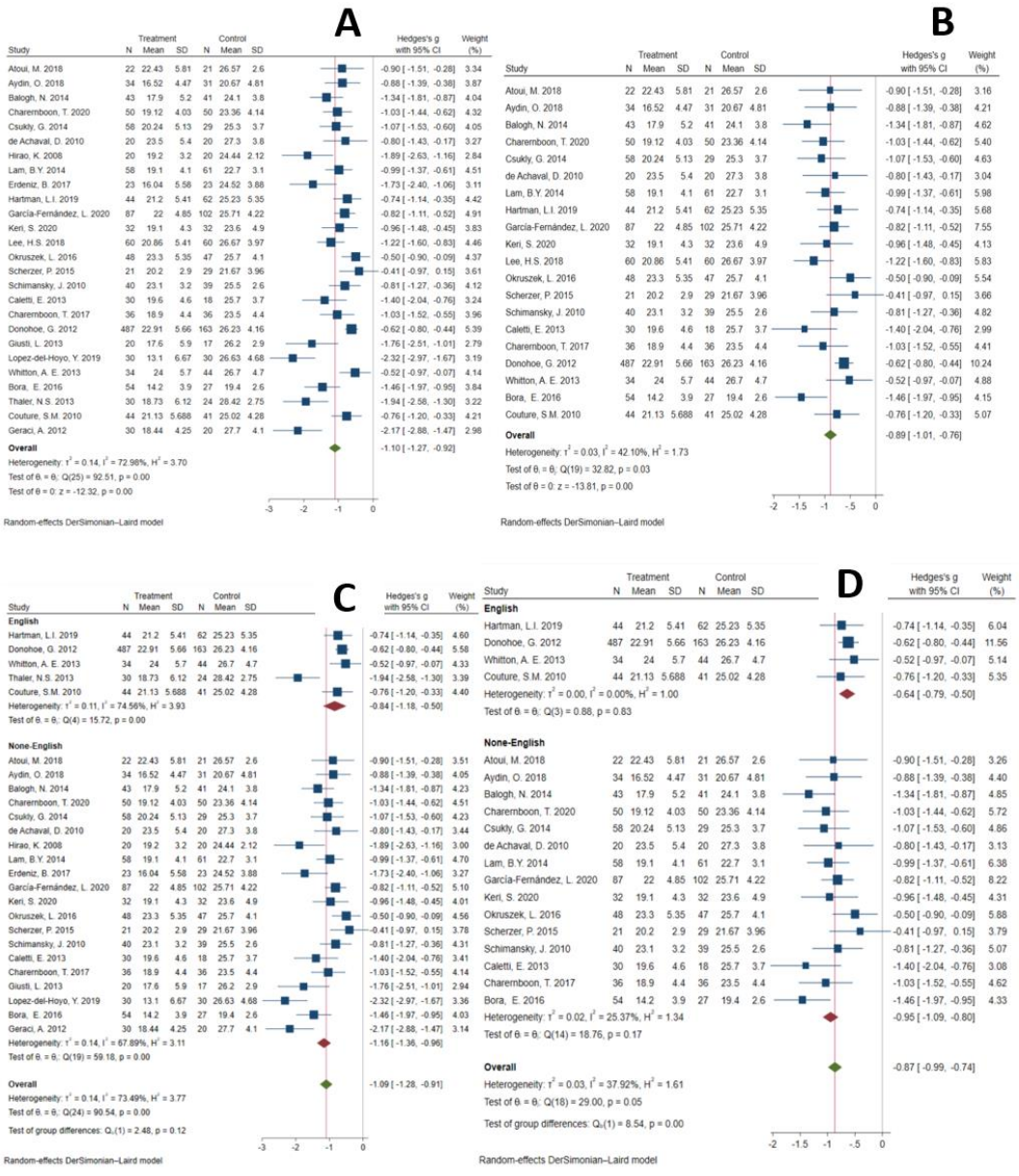

Figure 3.4 Forest plot of the standardized mean difference of RMET score between different types of respondents

Panel A: comparison of individuals with schizophrenia and healthy controls (26 studies)

Panel B: comparison of individuals with schizophrenia and healthy controls after removing the outliners (20 studies)

Panel C: comparison of individuals with schizophrenia and healthy controls stratified by the version of RMET (26 studies)

Panel D comparison of individuals with schizophrenia and healthy controls stratified by the version of RMET after removing the outliners (20 studies)

There was substantial heterogeneity in the estimated effect sizes of the 26 studies: the I^2 value was 73.0%, and the corresponding Q-statistic value was 92.5 ($p<0.001$). The funnel plot for the 26 studies (Figure 3.5A) identifies the main reason for this heterogeneity – the six smallest studies (total sample sizes ranging from 37 to 60) have the six largest effect sizes. Thus, the potential for publication bias is high, a finding supported by the results of Egger's test ($z=-4.53$, $p<0.001$). After removing these six outliers, the funnel plot for the remaining 20 studies is balanced (Figure 3.5B); the pooled standardized mean difference is reduced but still statistically significant (SMD=-0.89; $z=-13.81$, $p<0.001$); moreover, the I^2 value is reduced to 42.1% and the corresponding Q-test value was 32.8 ($p=0.03$) (Figure 3.4B).

Among the 26 studies, five studies used the original English version of RMET (Baron-Cohen, 2001), one study administered the English version to half of the participants and a Korean version to the other half, and 20 studies used translated versions of RMET (Turkish, Hungarian, Italian, and Spanish were each used in three papers; Thai was used in two papers; and Chinese, French, German, Japanese, Lebanese, and Polish were each used in a single paper). Based on the stratified analyses (Figure 3.4C), the pooled SMD was greater in the 20 studies using non-English versions (SMD=-1.16, $z=11.22$, $p<0.001$) than that in the five studies using the English version (SMD=-0.84, $z=3.28$, $p=0.001$), and heterogeneity was slightly greater in studies using the English version ($I^2=74.6\%$, $p<0.001$) than in studies using non-English versions ($I^2=67.9\%$, $p<0.001$). The SMD was not significantly different between these language-based subgroups when all 25 study samples were included in the analysis (Chi[Q]=2.48, $p=0.12$). However, after excluding the six outlier studies (Figure 4D), the SMD in the remaining 15 non-English RMET studies was significantly greater than the SMD in the remaining four English RMET studies (-0.95 vs -0.64,

Chi[Q]=8.54, p<0.001), but the four remaining studies using the English version were less heterogeneous than the 15 remaining studies that used non-English versions (I^2=0.0% in the four English RMET studies and I^2=37.9% in the 15 non-English RMET studies).

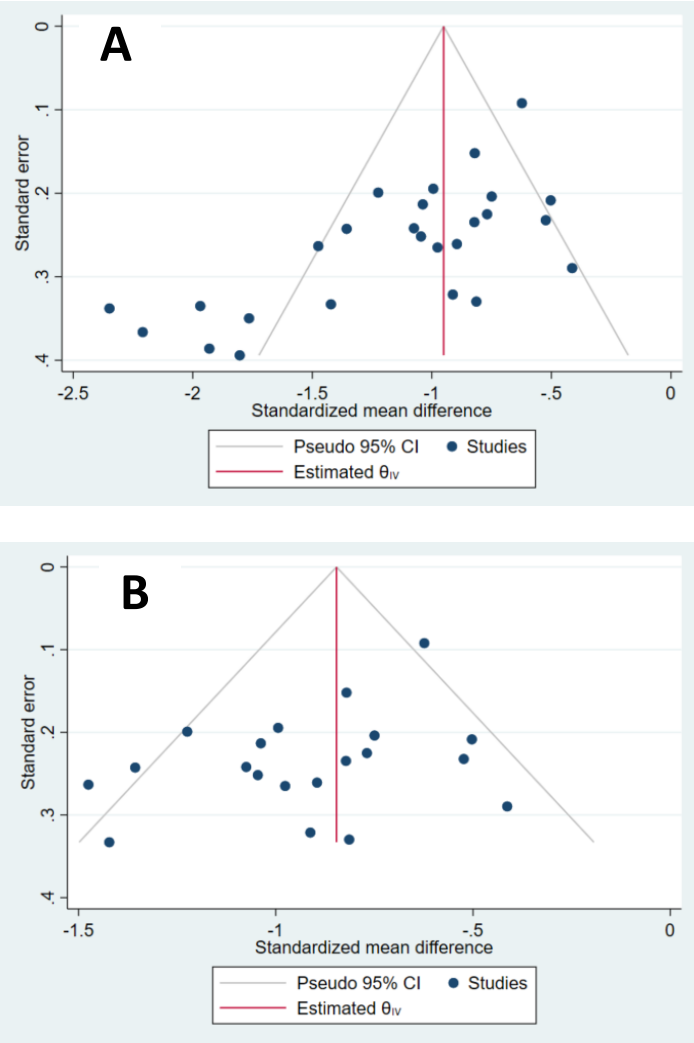

Figure 3.5 Funnel plots of results of meta-analyses

Panel A: Results of all 26 studies comparing individuals with schizophrenia and healthy controls

Panel B: Results of 20 studies that remain after removing studies with outlier results

3.3.5 Meta-regression on the covariates

There were 36 studies with 40 distinct samples of individuals with schizophrenia (combined sample size=1823) that provided both the mean age of the sample and the mean and standard deviation of the RMET scores; 29 of these studies included 35 distinct samples with schizophrenia (combined sample size=1620) that also provided the mean years of schooling of the sample. These data made it possible to conduct three separate regression analyses that included age, schooling, and both age and schooling as independent variables. Each regression equation was estimated using two methods: restricted maximum likelihood and the bootstrap method. As shown in Table 3.3, neither age, years or schooling nor the combination of age and years of schooling showed significant effects on RMET performance in the patient samples. when the regression only had age as an independent variable (Model 1, Figure 3.6A), the RMET score decreased with increasing age, but this decreasing trend was not statistically significant (β=-0.045, p=0.516). When the regression only included years of schooling as an independent variable (Model 2, Figure 3.6B), the RMET score increased with increasing years of schooling, but this increasing trend was not statistically significant (β=0.399, p=0.149).

Multivariate meta-regression using both mean age and mean years of schooling as independent variables (Model 3) also showed the negative relationship between RMET score and age (β=-0.032, p=0.635) and the positive relationship between RMET score and years of schooling (β=0.418, p=0.140) in patients with schizophrenia, but neither of these associations were statistically significant. The results using the two estimation methods were quite similar, but the p-values for the coefficients related to years of schooling are substantially smaller when using the bootstrap method.

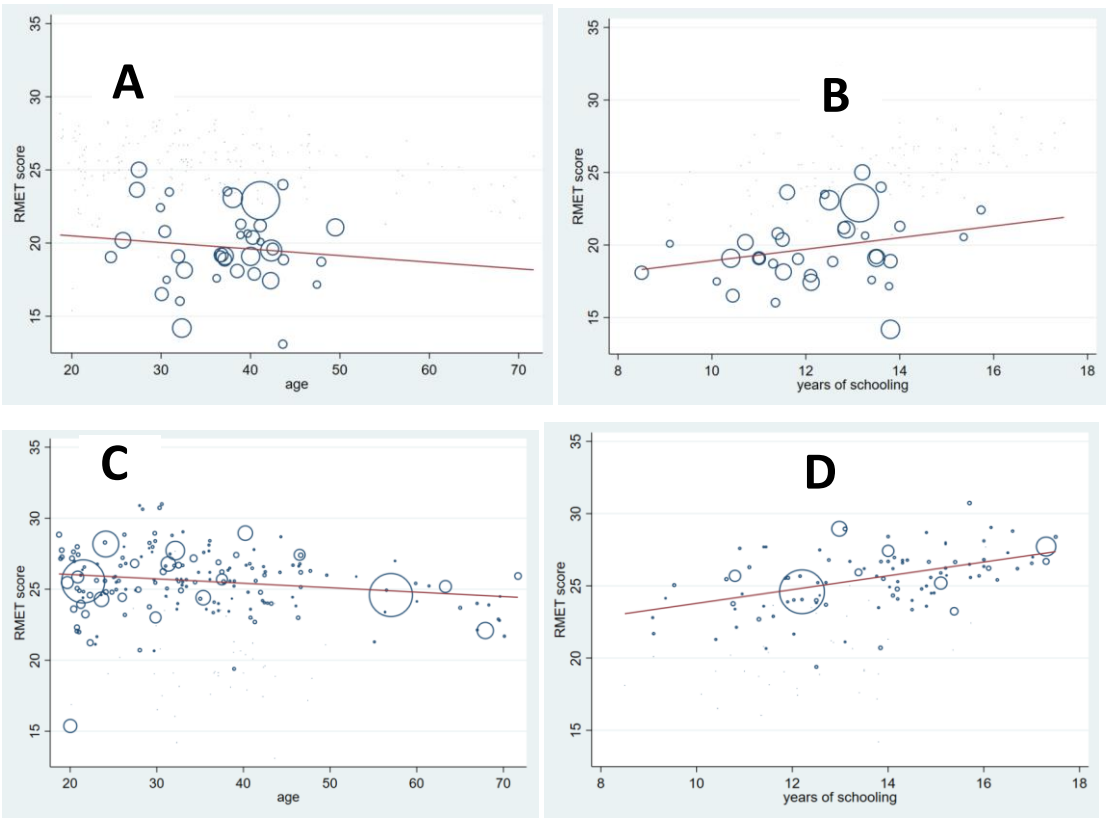

Figure 3.6 Association of age and years of schooling with RMET score in different respondents based on univariate meta-regression

Panel A: Association of age and RMET score in individuals with schizophrenia in 40 study samples

Panel B: Association of years of schooling and RMET score in individuals with schizophrenia in 35 study samples

Panel C: Association of age and RMET score in healthy controls in 180 study samples

Panel D: Association of years of schooling and RMET score in healthy controls in 99 study samples

Table 3.3 Meta-regression of RMET score on age and years and schooling

	Covariates	Individuals with schizophrenia			Healthy controls		
		Number of samples	Coef	P value	Number of samples	Coef	P value
Model 1	Age	40	-0.045	0.516	180	-0.031	**0.020**
Model 1 using bootstrap			-0.045	0.527		-0.031	**0.018**
Model 2	Years of schooling	35	0.399	0.149	98	0.477	**<0.001**
Model 2 using bootstrap			0.399	0.076		0.477	**<0.001**

Model 3	Age	35	-0.032	0.635	99	-0.026	0.126
	Years of schooling		0.418	0.140		0.423	**<0.001**
	Constant		15.88	**<0.001**		20.79	**<0.001**
Model 3 using bootstrap	Age		-0.032	0.648		-0.026	0.106
	Years of schooling		0.418	0.081		0.423	**<0.001**
	Constant		15.88	**<0.001**		20.80	**<0.001**

A parallel meta-regression analysis of healthy control subjects used the results from 180 distinct samples (combined sample size=21,494) that included data on the mean age of respondents; 98 of these samples (combined sample size=7946) also included data on the mean years of schooling of respondents. In these analyses, the regression that only included age as an independent variable (Model 1, Figure 3.6C) identified a statistically significant decrease in RMET scores with increasing age (β=-0.031, p=0.020); the regression that only included years of schooling as an independent variable (Model 2, Figure 3.6D) found a statistically significant increase in RMET scores with increasing years of schooling (β=0.477, p<0.001); and the multivariate meta-regression that included both age and years of schooling as independent variables (Model 3) found that increasing years of schooling remained significantly associated with increasing RMET scores (β=0.423, p<0.001) but the relationship of increasing age with decreasing RMET scores was no longer statistically significant (β=-0.026, p=0.126). In this case, the only difference between the two estimation methods was a smaller p-value for age in Model 3.

The differences in the association of age and education with RMET scores between the patient samples and healthy control samples may be related to the number of distinct samples available for the different analyses. For example, in the regressions using age as an independent variable, the coefficient for the 40 patient samples was substantially greater than that for the 180 healthy control samples (β=-0.045 vs β=-0.031), but the relationship of decreasing RMET scores with increasing age in the healthy control samples was statistically significant whereas that in the patient samples was not. Similarly, in the multivariate meta-regression analysis, the coefficient for the adjusted relationship of years of schooling in the 35 patient

samples (β=0.418) is essentially identical to that for the 99 healthy control samples (β=0.423), but the relationship of increasing RMET scores with increasing years of schooling is not statistically significant for the patient groups (p=0.140) while it is statistically significant for the healthy control groups (p<0.001).

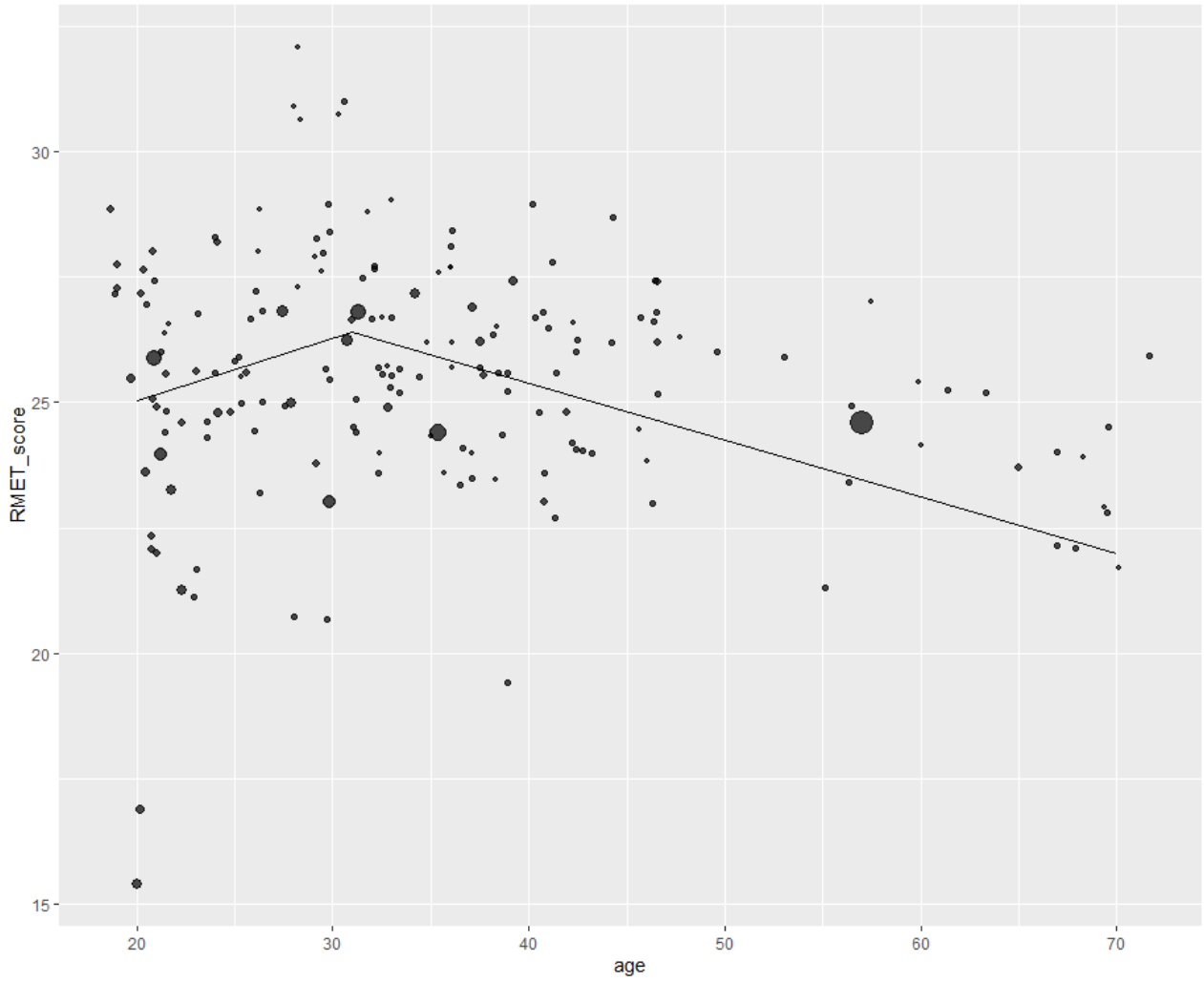

Figure 3.7 Relationship between age and RMET score in healthy controls using univariate linear regression with spline construction in 180 separate study samples.

In the multivariate meta-regression, the larger negative coefficient for age in the patient samples compared to that in the healthy control samples (β=-0.032 vs β=-0.026) suggests that after adjusting for years of schooling, the annual rate of decline in social cognitive functioning (as assessed by RMET) in patients with schizophrenia is

23% ([0.032-0.026]/0.026) faster than that in healthy controls.

3.3.6 Assessment of non-monotonic relationship between age and RMET score in healthy controls

The mean age of individuals in the 180 samples of healthy controls that included data on age ranged from 18.7 to 71.7 years of age, making it possible to assess a potential non-linear relationship of age with RMET scores using linear regression with spline construction. Assessing potential knots from 25 to 45 years of age, we identified 31 years of age as the point of inflexion (i.e., the knot with the lowest AIC) in both the univariate regression (only including age, AIC=792.2) and the multivariate analysis (including age and years of schooling, AIC=422.7). As shown in Table 3.4 and Figure 3.7, in the univariate analysis RMET score increased with age before age 31 (β=0.123, p=0.008) and declined with age after age 32 (β=-0.074, p<0.001). In the multivariate model (Table 3.4), after adjusting for years of schooling (which was significantly associated with RMET score), the RMET showed a significant increase with age before age 31 (β=0.179, p=0.048) and a statistically significant decline with age after age 31 (β=-0.048, p=0.011).

Table 3.4 Relationship of age and RMET score among healthy controls using univariate and multivariate meta-regression with spine construction

Covariates	Spine cut-off	Coefficient	P value
Age only (number of samples=180)	$\frac{dRMET}{dage}$|age⩽31	0.123	**0.008**
	$\frac{dRMET}{dage}$|age>31	-0.074	**<0.001**
	constant	22.59	**<0.001**
Age and years of schooling (number of samples=99)	$\frac{dRMET}{dage}$|age⩽31	0.179	**0.048**
	$\frac{dRMET}{dage}$|age>31	-0.048	**0.011**
	Years of schooling	0.427	**<0.001**
	constant	14.79	**<0.001**

3.4 Discussion

This review identified 198 studies that used RMET to assess social cognition in 41 separate samples of patients with schizophrenia and 197 separate samples of healthy controls. The pooled mean RMET score of the 1823 patients and 23,619 healthy controls included in these studies was much lower in patients than in healthy controls (19.8 [18.9-20.6] vs 25.5 [25.2-25.9], z=12.41, p<0.001). Meta-analysis of the results of 26 studies that directly compared RMET scores in patients with schizophrenia and healthy controls found that the pooled mean of patients' scores was more than one SMD lower than the pooled mean score of healthy controls. Significant publication bias was identified among these studies (studies with smaller sample sizes were more likely to report larger SMD between the two groups), but the differences between groups remained significant after removing the outlier studies (SMD=-0.89; z=-13.81, p<0.001). These results confirming previous findings that patients with schizophrenia suffer from substantial deficits in theory of mind.

The results for both patients and healthy controls were quite heterogenous, so we used meta-regression methods to explore the relationship between mean RMET performance, mean age, and mean level of education in patient samples and, separately, in healthy control samples. In the univariate analyses, age was negatively related to RMET scores and educational level was positively related to RMET scores in both the patient samples and the healthy control samples, but the results were only statistically significant for the healthy control samples, possibly because of the much smaller number of patient samples available for analysis. A separate meta-regression with spline construction in the healthy control samples found that RMET scores increased with age before age 31 and decreased with age after age 31. (The much smaller number of samples of patients with schizophrenia and the smaller range in the mean age of these samples made it infeasible to conduct a spline construction meta-regression using the patient samples.) These relationships persisted in the multivariate analysis (that included both age and years of schooling as covariates), though the effect of age was attenuated after adjustment for years of schooling.

Previous findings about the relationship of age with RMET scores have been inconsistent. Dodell-Feder and colleagues (2020) used online interviews to assess RMET in 40,248 participants 10-70 years of age and found that RMET scores

increased with age up until age 65. Cabinio and colleagues (2015) reported *unchanging* RMET scores in healthy respondents 20-70 years of age. Two cross-sectional studies (Slessor, G., 2007; Javkowiak-Siuda, K, 2016) comparing RMET performance in persons over 65 to that of persons under 35 found that the older participants had significantly *lower* RMET scores. Finally, Pardini (2009), Deng (2021), and Lee (2020) reported that RMET performance started to decline in the fifth decade of life, at age 60, and at age 66, respectively.

Our systematic review of 198 studies that administered RMET to 180 separate samples of healthy subjects is the first study to identify a non-monotonic relationship between RMET score and age, suggesting that individuals accumulate knowledge and skills of theory of mind until they reach early middle age (31 years of age) and then their theory of mind skills gradually decline with normal aging. Factors that have been proposed as contributors to age-related declines in theory of mind may also contribute to this novel empirical finding; these factors include impaired ability to decode cues (Slessor, G., 2007), declining executive function and processing speed (Charlton, R.A., 2009), impaired destination memory (El Haj, M., 2016), and selective cortical declines in gray matter (Cabinio, M., 2015).

Several hypotheses have been proposed to explain increasing deficits in theory of mind with aging. Slessor and colleagues (2007) suggested that deficits in theory of mind are manifestations of general impairment in the ability to decode cues. Some researchers suggest that decline of theory of mind is mediated by impairment in other cognitive domains, such as executive function, information processing speed (Charlton, R.A., 2009), destination memory (El Haj, M., 2016), and verbal intelligence (Slessor, G., 2007). And neuroimaging studies report that decline in RMET score with aging is correlated with decreasing volume in the bilateral precentral gyrus, bilateral posterior insula, left superior temporal gyrus, and left inferior frontal gyrus (Cabinio, M., 2015). This raises the possibility that the neurodevelopmental trajectory of this component of cognition is more prolonged than that of other types of cognition and, thus, can be disrupted at later ages by serious mental illnesses like schizophrenia.

The association of years of schooling with RMET scores was stronger than the association of age with RMET scores in the current findings, but there has been much

less research about the role of education in the development of theory of mind. Khorashad and colleagues (2018) found no significant relationship between RMET score and educational attainment, while other studies (Schmidt, 2009; Dodell-Feder, 2020; Deng, 2021) find that years of schooling can explain some variance in the RMET score.

Subgroup analyses indicated that after excluding the outlier studies the difference in RMET performance between patients with schizophrenia and healthy controls was greater in studies using non-English versions of RMET than in those using the original English version (Chi [Q]=8.54, $p<0.001$). The reason for this difference is unclear. All of the studies used the same sets of pictures (with Caucasian subjects), so it is likely (though not certain) that respondents administered non-English versions of RMET were *less* ethnically similar to the ethnic groups in the stimulus pictures than respondents administered the English version of RMET. The difficulty patients have in identifying emotions in the RMET may be magnified when presented with pictures of persons with an ethnicity different from their own, resulting in a greater assessed deficit compared to healthy controls in studies that use non-English versions of RMET. Further research with non-Caucasian pictures is needed to clarify this issue.

Related to the demographic differences in RMET performance, familiarity with the four verbal descriptors provided as potential response choices for each presented picture in the RMET is, presumably, a prerequisite for making the correct selection. It is reasonable to expect that persons with lower levels of education will have less familiarity with the presented terms, lower verbal cognitive test scores and, thus, greater difficulty achieving high RMET scores. Moreover, the relative difficulty of the terms associated with each picture and the distinctiveness of the meanings of the four presented terms will vary across languages, so it is likely that the association of education level with the total RMET score (and with the pattern of incorrect items) will vary for different language versions of the RMET. Assessment of item-difficultly in each language (e.g., their frequency of use in daily speech) and comparison of RMET scores with measures of verbal intelligence will be needed to 1) decide on the minimum education level appropriate for administering the RMET; 2) develop a method of adjusting RMET scores based on education level or vocabulary skill, and 3)

develop alternative versions of RMET suitable for persons with little formal education.

3.5 Limitations

There are several potential limitations. 1) We only searched for studies published in English or Chinese, so the analyses did not include studies published in other languages. 2) Some samples in the papers did not include data about key variables needed in the analysis (i.e., standard deviation of mean RMET score, age or educational level of sample), and some other studies were of low methodological quality. 3) Only 26 of the 198 studies directly compared RMET results of patients with schizophrenia and healthy controls, limiting our ability to conduct meta-analyses of results. 4) Most samples of patients with schizophrenia were chronic patients regularly using antipsychotic medications, so their deficits in theory of mind may not be representative of that in all individuals with schizophrenia. 5) The range in the mean age and mean years of education of the 40 samples of patients was relatively narrow, making it difficult to accurately assess the potential relation of age and education with RMET scores in the patients. 6) The distribution of the mean age of the 180 separate samples of healthy controls was imbalanced (the mean age of 88% of the samples was below 50), which potentially biased the assessment of the inflexion point (at 31 years of age) in the meta-regression spline construction analysis. 7) Few studies measured the vocabulary level of participants, making it impossible to explore the potential relationship of verbal ability with RMET performance.

3.6 Conclusion

This is the first systematic review and meta-analysis of studies using the Reading the Mind in the Eyes Test (RMET) to assess social cognitive functioning among individuals with schizophrenia. Meta-analyses of data from 198 studies that use the Reading the Mind in the Eyes Test (RMET) to assess theory of mind confirm previous findings that patients with schizophrenia experience severe impairments in theory of mind. RMET scores decrease with age and increase with years of schooling in both

patients and healthy controls, though these relationships were only statistically significant in the healthy control samples (possibly due to the much smaller number of patient samples available for analysis). In the multivariate meta-analysis of healthy control samples that included both age and years of schooling as covariates, years of schooling remained significantly associated with RMET scores, but age was no longer significantly associated with RMET scores. We also found a previously unreported non-monotonic relationship between age and RMET performance in healthy controls: the RMET score increased with age before age 32 and decreased with age after age 31. These findings highlight the need to clarify the relationships between education, verbal intelligence, and social cognition; they also suggest the need for a more nuanced assessment of the neurodevelopment of theory of mind—which may be different from the neurodevelopment of other cognitive abilities.

Funding

This study was supported by National Institute of Mental Health grants for the project 'Characterizing Cognition Across the Lifespan in Untreated Psychosis in China' (PI's: Yang, Phillips, Keshavan; number, MH108385 R01) and the project 'Longitudinal trajectories of treated and untreated psychosis (LoTUS)' (PI's: Yang, Phillips, Stone, Keshavan; number, R01 MH127631): the Shanghai Mental Health Qi Hang Project (PI, Deng; number, 2018-QH-06): and the Shanghai Mental Health Project (PI, Deng; number, 2021-YJG05). None of these funding agencies had any role in the design or conduct of the study; in the collection, management, analysis, or interpretation of the data; in the preparation, review, or approval of the manuscript; or in the decision to submit the manuscript for publication.

Contributors

FD designed the study, coordinated the collection and analysis of the data, and wrote the initial draft of the manuscript. MAB, YRC, JT, XB, YC, JL, ZL, and QY screened articles and extracted data from selected articles. MQ provided advice about data analysis. LHY made revisions to the initial draft. MRP, LHY, and WSS provided technical support throughout all the steps of the study. MRP made detailed revisions to the manuscript.

Declaration of interests

The authors report no conflict of interest related to this manuscript.

References

Achim, A. M., Ouellet, R., Roy, M. A., & Jackson, P. L. (2012). Mentalizing in first-episode psychosis. *Psychiatry Research*, 196, 207–213. doi:10.1016/ j.psychres.2011.10.011.

Adenzato, M., Todisco, P., & Ardito, R.B. (2012). Social Cogntion in Anorexia Nervosa: Evidence of Preserved Theory of Mind and Impaired Emotional Functioning. Plos One, 7(8), e44414.

Adenzato, M., Manenti, R., Enrici, I., Gobbi, E., Brambilla, M. Alberici, A., … Cotelli, M. (2019). Transcranial direct current stimulation enhances theory of mind in Parkinson's disease patients with mild cognitive impairment: a randomized, double-blind, sham-controlled study. Translational Neurodegeneration, 8, 1. Doi: doi.org/10.1186/s40035-018-0141-9

Ahmed, S.A. & Miller, L.S. (2011). Executive Function Mechanisms of Theory of Mind. *J Austism Dev Disord*, 41, 667-678

Ainley, V., Maister, L., & Tsakiris, M. (2015). Hearthelf empathy? No association between interoceptive awareness, questionnaire measures of empathy, reading the mind in the eyes task or director task. *Frontiers in Psychology,* 6, 554. Doi: 10.3389/fpsyg.2015.00554

Altunbas, F.D., Unsalver, B.O., & Yasar, A.B. (2019). Aspects of cognitive performance relating to Theory of Mind (ToM) among people diagnosed with Post-Traumatic Stress Disorder (PTSD). *Neuropsychiatric Disease and Treatment*, 15, 2015-2025.

Anupama, V., Poornima, B., Jagadisha, T., & Urvakhsh, M.M. (2018). Pattern of Social Cognition Deficits in Individuals with Borderline Personality Disorder. *Asian Journal of Psychiatry*, 33, 105-112. Doi: 10.1016/j.ajp.2018.03.010

Atoui, M., Jamil, F.E., Khoury, J.E., Doumit, M., Syriani, N., Khani, M., & Nahas, Z. (2018). The relationship between clinical insight and cognitive and affective empathy in schizophrenia. *Schizophrenia Research Cognition*, 12: 56-65. Doi: 10.1016/j.scog.2018.02.004

Aydin, O., Lysaker, P.H., Balıçı, K., Ünal-Aydin, P., & Esen-Danaci, A. (2018). Associations of oxytocin and vasopressin plasma levels with neurocognitive, social cognitive and meta cognitive function in schizophrenia. *Psychiatry Research*, 270, 1010-1016. Doi: doi.org/10.1016/j.psychres.2018.03.048

Ayesa-Arriola, R., Setién-Suero, E., Neergaard, K.D., Ferro, A., Fatjó-Vilas, M., Ríos-Lago, M., … Crespo-Facorro, B. (2016). Evidence for Trait Related Theory of Mind Impairment in First Episode Psychosis Patients and Its Relationship with Processing Speed: A 3 Year Follow-up Study. *Frontiers in Psychology*, 7, 592. Doi: 10.3389/fpsyg.2016.00592

Bai, Q.Y. (2011). Adolescent Theory of Mind Development and its Affected Personality Factors: Machiavellian Beliefs. [Master dissertation, Southwest University] [in Chinese]

Balogh, N., Égerházi, A., Berecz, R., & Csukly, G. (2014). Investigating the state-like and trait-like characters of social cognition in schizophrenia: A short term follow-up study. *Schizophrenia Research*, 159, 499-505. Doi: dx.doi.org/10.1016/j.schres.2014.08.027

Banati, M., Sandor, J., Mike, A., Illes, E., Bors, L., Feldmann, R., … Illes, Z. (2009). Social cognition and Theory of Mind in patients with relapsing-remitting multiple sclerosis. *European Journal of Neurology*, 17, 426-433. Doi: 10.1111/j.1468-1331.2009.02836.x

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The Reading the Mind in the Eyes Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 42(2), 241–251. doi10.1111/1469-7610.00715.

Baron-Cohen, S., Bowen, D.C., Holt, R.J., Allison, C., Anyeung, B., Lombardo, M.V., … Lai, M.C. (2015). The “Reading the Mind in the Eyes” Test: Complete Absence of Typical Sex Difference in ~400 Men and Women with Autism. *PLOS One*, 10(8), e0136521. Doi: 10.1371/journal.pone.0136521

Barrera, Á., Vázquez, G., Tannenhaus, L., Lolich, M, & Herbst, L. (2013). Theory of mind and functionality in bipolar patients with symptomatic remission. *Revista de Psiquiatría y Salud Mental*, 6(2), 67-74.

Bayliss, L., Galvez, V., Ochoa-Morales, A., Chávez-oliveros, M., Rodríguez-Agudelo, Y., Delgado-García, G., & Boll, M.C. (2019). Theory of mind impairment in Huntington's disease patients and their relatives. *Arquivos de Neuro-Psiquiatria*, 77(8), 574-578. doi: https://doi.org/10.1590/0004-282X20190092

Bedwell, J.S., Compton, M.T., Jentsch, F.G., Deptula, A.E., Goulding, S.M, & Tone, E.B. (2014). Latent Factor Modeling of Four Schizotypy Dimensions with Theory of Mind and Empathy. *Plos One*, 9(11), e113853. Doi: 10.1371/journal.pone.0113853

Berenson, K.R., Dochat, C., Martin, C.G., Yang, X., Rafeali, E., & Downey, G. (2018). Identification of mental states and interpersonal functioning in borderline personality disorder. *Personal Disord*, 9(2), 172-181. Doi: 10.1037/per0000228

Besnard, J., Gall, D.L., Chauviré, V., Aubin, G., Etcharry-Bouyx, F., & Allain, P. (2016). Discrepancy between social and non-social decision-making under uncertainty following prefrontal lobe damage: The impact of an interactionist approach. *Social Neuroscience*, 4, 430-447. Doi: doi.org/10.1080/17470919.2016.1182066

Black, J.E. (2019). An IRT Analysis of the Reading the Mind in the Eyes Test. *Journal of Personality Assessment*, 101, 425-433. Doi: doi.org/10.1080/00223891.2018.1447946

Bodnar, A., & Rybakowski, J.K. (2017). Mentalization deficit in bipolar patients during an acute depressive and manic episode: association with cognitive functions. *International Journal of Bipolar Disorders*, 5, 38. Doi: 10.1186/s40345-017-0107-3.

Bora, E., Sehitoglu, G., Aslier, M., Atabay, I., & Veznedaroglu, B. (2007). Theory of mind and unawareness of illness in schizophrenia. *Eur Arch Psychiatry Clin Neurosci*, 257, 104-111.

Bora, E., Veznedzroglu, B., & Vahip, S. (2016). Theory of mind and executive functions in schizophrenia and bipolar disorder: A cross-diagnostic latent class analysis for identification of neuropsychological subtypes. *Schizophrenia Research*, 176, 500-505. Doi: dx.doi.org/10.1016/j.schres.2016.06.007

Boucher, O., Rouleau, I., Lassonde, M., Lepore, F., Bouthillier, A., & Nguyen, D.K. (2015). Social information processing following resection of the insular cortex. *Neuropsychologia*, 71, 1-10. Doi: dx.doi.org/10.1016/j.neuropsychologia.2015.03.008

Brambilla, M., Cotelli, M., Manenti, R., Dagani, J., Sisti, D., Rocchi, M., … Girolamo, G.D. (2016). Oxytocin to modulate emotional processing in schizophrenia: A randomized, double-blind, cross-over clinical trial. *European Neuropsychopharmacology*, 26(10): 1619-1628

Budak, E.A., Küçükgöncü, S., & Bestepe, E.E. (2019). Theory of mind abilities and its relationship with clinical features in euthymic bipolar patients. *Cukurova Medical Journal*, 44(1), 176-182. Doi: 10.17826/cumj.497219

Buhlmann, U., Winter, A., & Kathmann, N. (2013). Emotion recognition in body dysmorphic disorder: Application of the Reading the Mind in the Eyes Task. *Body Image*, 10, 247-250

Burke, T., Pinto-Grau, M., Lonergan, K., Elamin, M., Bede, P., Costello, E., … Pender, N. (2016). Measurement of Social Cognition in Amyotrophic Latent Sclerosis: A Population Based Study. *Plos One*, 11(8), e016850. Doi: 10.1371/journal.pone.0160850

Cabinio, M., Rossetto, F., Blasi, V., Savazzi, F., Castelli, I., Massaro, D., … Baglio, F. (2015). Mind-Reading Ability and Structural Connectivity Changes in Aging. *Frontiers in Psychology*, 6, 1808. Doi: 10.3389/fpsyg.2015.01808

Caletti, E., Paoli, R.A., Fiorentini, A., Cigliobianco, M., Zugno, E., Serati, M., … Altamura, A.C. (2013) Neuropsychology, social cognition and global functioning among bipolar, schizophrenic patients and healthy controls: preliminary data. *Frontiers in Human Neuroscience*, 7:1-14. Doi: 10.3389/fnhum.2013.00661

Carré, J.M., Ortiz, T.L., Labine, B., Moreau, B.J.P, Viding, E., Neumann, C.S, & Goldfarb, B. (2015). Digit ratio (2D:4D) and psychopathic traits moderate the effect of exogenous testosterone on socio-cognitive processes in men. *Psychoneuroendocrinology*, 62, 319-326. Doi: dx.doi.org/10.1016/j.psyneuen.2015.08.023

Cavallo, M., Adenzato, M., MacPherson, S.E., Karwig, G., Enrici, I, & Abrahams, S. (2011). Evidence of Social Understanding Impairment in Patients with Amyotrophic Lateral Sclerosis. *Plos One*, 6(10), e25948. Doi: 10.1371/journal.pone.0025948

Charernboon, T., & Patumanond, J. (2017). Social cognition in schizophrenia. *Mental Illness*, 9, 7054. Doi: 10.4081/mi.2017.7054

Charlton, R.A, Barrick, T.R., Markus, H.S, & Morris, R.G. (2009). Theory of mind associations with other cognitive functions and brain imaging in normal aging. *Comparative Study*, 24(2), 338-48. Doi: 10.1037/a0015225

Chen, K.W., Lee, S.C., Chiang, H.Y., Syu, Y.C., Yu, X.X., & Hsieh, C.L. (2017). Psychometric properties of three measures assessing advanced theory of mind: Evidence from people with schizophrenia. *Psychiatry Research*, 257, 490-496. Doi: 10.1016/j.psychres.2017.08.026

Coyle, T.R., Elpers, K.E., Gonzalez, M.C., Freeman, J., & Baggio, J.A. (2018). General Intelligence(g), ACT Scores, and Theory of Mind: (ACT)g Predicts Limited Variance Among Theory of Mind Tests. *Intelligence*, 71, 85-91. Doi: doi.org/10.1016/j.intell.2018.10.006

Craig, J.S. (2004). Persecutory beliefs, attributions and theory of mind: comparison of patients with paranoid delusions, Asperger's syndrome and healthy controls. *Schizophrenia Research*, 69, 29-33.

Csukly, G., Polgár, P., Tombor, L., Benkovits, J., & Réthelyi, J. (2014). Theory of mind impairments in patients with deficit schizophrenia. *Comprehensive Psychiatry*, 55, 349-356. Doi: dx.doi.org/10.1016/j.comppsych.2013.08.025

Cusi, A.M., MacQueen, G.M., & Mckinnon, M.C. (2012). Patients with bipolar disorder show impaired performance on complex tests of social cognition. *Psychiatry Research*, 200, 258-264.

Corcoran, R., Mercer, G., & Frith, C. D. (1995). Schizophrenia, symptomatology and social inference: Investigating “theory of mind” in people with schizophrenia. *Schizophrenia Research*, 17(1), 5–13. Doi: 10.1016/ 0920-9964(95)00024-G.

Cotelli, M., Manenti, R., Gobbi, E., Enrici, I., Rusich, D., Ferrari, C., & Adenzato, M. (2020). Theory of Mind Performance Predicts tDCS-Medicated Effects on the Medial Prefrontal Cortex: A Pilot Study to Investigate the Role of Sex and Age. *Brain Science*, 10, 257. Doi: 10.3390/brainsci10050257

Coutture, S.M., Penn, D.L., Losh, M., Adolphs, R., Hurley, R., & Piven, J. (2010). Comparison of social cognitive functioning in schizophrenia and high functioning autism: more convergence than divergence. *Comparative Study*, 40(4), 569-579.

Dal Monte, O., Schintu, S., Pardini, M., Berti, A., Wassermann, E.M., Grafman, J., & Krueger, F. (2014). The left inferior frontal gyrus is crucial for reading the mind in the eyes: Brain lesion evidence. *Cortex*, 58, 9-17. Doi: dx.doi.org/10.1016/j.cortex.2014.05.002

Dalkner, N., Bengesser, S.A., Birner, A., Fellendorf, F.T., Hamm, C., Platzer, M., … Reininghaus, E.Z. (2019). The relationship between “Eyes Reading” ability and verbal memory in bipolar disorder. *Psychiatry Research*, 273, 42-51. Doi: doi.org/10.1016/j.psychres.2019.01.015

de Achával, D., Costanzo, E.Y., Villarreal, M., Jáuregui, I.O., Chiodi, A., Mariana, N., … Guinjoan, S.M. (2010). Emotion processing and theory of mind in schizophrenia patients and their unaffected first-degree relatives. *Neuropsychologia*, 48, 1209-1215. Doi: 10.1016/j.neuropsychologia.2009.12.019

De Los Reyes, A., Lerner, M.D., Thomas, S.A., Daruwala, S., & Goepel, K. (2013). Discrepancies between Parent and Adolescent Beliefs about Daily Life Topics

and Performance on an Emotion Recognition Task. *J Abnorm Child Psychol*, 41(6), 971-982. Doi: 10.1007/s10802-013-9733-0.

De Sampaio, F.T., Soneira, S., Aulicino, A., & Allegri, R.F. (2013). Theory of Mind in Eating Disorders and Their Relationship to Clinical Profile. *Eur Eat Disorders Rev*, 21, 479-487. Doi: 10.1002/erv.2247

DeAgnelo, G., & McCannon, B.C. (2017). Theory of Mind predicts cooperative behavior. *Economics Letters*, 155, 1-4. Doi: dx.doi.org/10.1016/j.econlet.2017.02.009

Dehning, S., Gasperi, S., Krause, D., Meyer, S., Reiss, E., Burger, M, … Siebeck, M. (2013). Emotional and Cognitive Empathy in First-Year Medical Students. *ISRN Psychiatry*, 2013, 801530. Doi: dx.doi.org/10.1155/2013/801530

Dehning, S., Gasperi, S., Tesfaye, M., Girma, E., Meyer, S., Krahl, W., … Siebeck, M. (2013). Empathy without borders? Cross-culture heart and mind-reading in first-year medical students. *Ethiop J Health Sci*, 23(2), 113-122.

Demartini, B., Petrochilos, P., Ricciardi, L., Price, G., Edwards, M.J., & Joyce, E. (2014). The role of alexithymia in the development of functional motor symptoms (conversion disorder). *Neuropsychiatry*, 85, 1132-1137. Doi: 10.1136/jnnp-2013-307203

Deng, F., Phillips, M.R., Cai, B., Yu, G., Qian., M, Grivel, M.M.R, … Yang, L.H. (2021). Comparison of social cognition using an adapted Chinese version of the Reading the Mind in the Eyes Test in drug-naïve and regularly medicated individuals with chronic schizophrenia and healthy controls in rural China. Psychological Medicine. Advance online publication. Doi: doi.org/10.1017/S003329172100043X.

DerSimonian R., & Laird N. (1985). Meta-analysis in clinical trials. *Control Clin Trials*, 7, 177-88.

Dodell-Feder, D., Ressler, K.J., & Germine, L.T. (2020). Social cognition or social class and culture? On the interpretation of differences in social cognitive performance. *Psychological Medicine*, 50(1): 133-145.

Domes, G., Heinrichs, M., Michel, A., Berger, C., & Herpertz, S. C. (2007). Oxytocin Improves “Mind-Reading” in Humans. *Biol Psychiatry*, 61, 731-733.

Donohoe, G., Duignan, A., Hargreaves, A., Morris, D.W., Rose, E., et al. (2012). Social Cognition in bipolar disorder versus schizophrenia: comparability in mental state decoding deficits. *Bipolor Disorder*, 14, 743-748.

Dulau, C., Deloire, M., Diaz, H., Saubusse, A., Charre-Morin, J., Prouteau, A, & Brochet, B. (2017). Social cognition according to cognitive impairment in different clinical phenotypes of multiple sclerosis. *J Neuro*, 264, 740-748. Doi: 10.1007/s00415-017-8417-z

Duque-Alarcón, X., Alcalá-Lozano, R., González-Olvera, J.J., Garza-Villarreal, E.A., & Pellicer, F. (2019). Effects of Childhood Maltreatment on Social Cognition and Brain Functional Connectivity in Borderline Personality Disorder Patients. *Frontiers in Psychiatry*, 10, 156. Doi: 10.3389/fpsyt.2019.00156

Eddy, C.M., & Rickards, H.E. (2015). Theory of Mind Can Be Impaired Prior to Motor Onset in Huntington’s Disease. *Neuropsychology*, 29(5), 792-798. Doi: dx.doi.org/10.1037/neu0000190

Eddy, C.M., Rickards, H.E., & Hansen, P.C. (2018). Through your eyes or mine? The neural correlates of mental state recognition in Huntington’s disease. *Human Brain Mapp*, 39, 1354-1366. Doi: 10.1002/hbm.23923

Eddy, C.M., & Hansen, P.C. (2020). Predictors of performance on the Reading the Mind in the Eyes Test. *Plos One*, 15(7), e0235529. Doi: doi.org/10.1371/journal.pone.0235529

Egger M., Davey S. G., Schneider M., Minder C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315, 629–34.

El Haj, M., Raffard, S., & Gély-Nargeot, M. (2016). Destination memory and cognitive theory of mind in normal ageing. *Memory*, 24(4), 526-534. Doi: dx.doi.org/10.1080/096

Enrici, I., Adenzato, M., Ardito, R.B., Mitkova, A., Cavallo, M., Zibetti, M., … Castelli, L. (2015). Emotion Processing in Parkinson’s Disease: A Three-Level Study on

Recognition, Representation, and Regulation. *Plos One*, 10(6), e013470. Doi: 10.1371/journal.pone.0131470

Enrici, I., Mitkova, A., Castelli, L., Lanotte, M, Lopiano, L., & Adenzato, M. (2017). Deep Brain Stimulation of the subthalamic nucleus does not negatively affect social cognitive abilities of patients with Parkinson's disease. *Scientific Reports*, 7, 9413. Doi: 10.1038/s41598-017-09737-6

Erdeniz, B., Serin, E., Íbadi, Y., & Tas, C. (2017). Decreased functional connectivity in schizophrenia: The relationship between social functioning, social cognition and graph theoretical network measures. *Psychiatry Research: Neuroimaging*, 270, 22-31. Doi: dx.doi.org/10.1016/j.psychresns.2017.09.011

Espinós, U., Fernández-Abascal, E.G., & Ovejero, M. (2018). What your eyes tell me: Theory of mind in bipolar disorder. *Psychiatry Research*, 262. 536-541. Doi: dx.doi.org/10.1016/j.psychres.2017.09.039

Fernández-Abascal, E.G, Cabello, R., Fernández-Berrocal, P., & Baron-Cohen, S. (2013). Test-retest reliability of the 'Reading the Mind in the Eyes' test: a one-year follow-up study. *Molecular Autism*, 4, 33.

Fertuck, E.A., Jekal, A., Song, I., Wyman, B., Morris, M.C., Wilson, S.T., … Stanley, B. (2009). Enhanced 'Reading the Mind in the Eyes' in borderline personality disorder compared to healthy controls. *Psychological Medicine*, 39(12), 1979-1988. doi:10.1017/S003329170900600X.

Filippo, S.G., Guida, M., Caciagli, L., Pagni, C., Pizzanelli, C., Bonanni, E., … Bonuccelli, U. (2016). Social cognition in Juvenile Myoclonic Epilepsy. *Epilepsy Research*, 128, 61-67. Doi: dx.doi.org/10.1016/j.eplepsyres.2016.10.017

Fossati, A., Borroni, S., Dziobek, I., Fonagy, P., & Somma, A. (2018). Thinking about assessment: Further evidence of the validity of the Movie for the Assessment of Social Cognition as a measure of mentalistic abilities. *Psychoanalytic Psychology, 35*(1), 127–141. https://doi.org/10.1037/pap0000130

García-Fernández, L., Cabot-Ivorra, N., Romero-Ferreiro, V., Pérez-Martín, J., & Rodriguez-Jimenez, R. (2020). Differences in theory of mind between early

and chronic stages in schizophrenia. *Journal of Psychiatric Research*, 127, 35-41. Doi: doi.org/10.1016/j.jpsychires.2020.05.009

Gavilá, J.M., & Haro, J. (2017). Personality traits and theory of mind: Performance data of a Spanish sample of university students. *Data in Brief*, 14, 612-617. Doi: dx.doi.org/10.1016/j.dib.2017.08.014

Geraci, A., Surian, L., Ferraro, M, & Cantagallo, A. (2010). Theory of Mind in patients with ventromedial or dorsolateral prefrontal lesions following traumatic brain injury. *Brain Injury*, 24, 7-8, 978-987. Doi: 10.3109/02699052.2010.487477

Geraci, A., Signorelli, M.S., & Aguglia, E. (2012). Reading the mind: a comparative study of out- and inpatients. *Journal of Psychopathology*, 18: 247-250.

Giordana, M., Licea-Haquet, G., Navarrete, E., Valles-Capetillo, E., Lizcano-Cortés, F., Carrillo-Peña, A. & Zamora-Ursulo, A. (2019). Comparison between the Short Story Task and the Reading the Mind in the Eyes for evaluating Theory of Mind: A replication report. *Cogent Psychology*, 6, 1, 1634326. Doi: 10.1080/23311908.2019.1634326

Giusti, L., Mazza, M., Pollice, R., Casacchia, M., & Roncone, R. (2013). Relationship between self-reflectivity, Theory of Mind, neurocognition, and global functioning: An investigation of schizophrenic disorder. *Clinical Psychologist*, 17, 67-76. Doi: 10.1111/cp.12006

Goddard, E., Carral-Fernández, L., Denneny, E., Campbell, I.C., & Treasure, J. (2014). Cognitive flexibility, central coherence and social emotional processing in males with an eating disorder. *The World Journal of Biological Psychiatry*, 15, 4, 317-326. Doi: 10.3109/15622975.2012.750014

Gong, P.Y., Liu, J.T., Li, S., & Zhou, X.L. (2014). Dopamine beta-hydroxylase gene modulates individuals' empathic ability. *SCAN*, 9, 1341-1345. Doi: 10.1093/scan/nst122

Gooding, D.C., Johnson, M. & Peterman, J.S. (2010). Schizotypy and altered digit ratios: a second look. *Psychiatry Research*, 178, 73-78. Doi: 10.1016/j.psychres.2010.04.023

Gooding, D.C., & Pflum, M.J. (2011). Theory of Mind and Psychometric Schizotypy. *Psychiatry Research*, 188: 217-223. Doi: 10.1016/j.psychres.2011.04.029

Grainger, S.A., Henry, J.D., Naughtin, C.K., Comino, M.S, & Dux, P.E. (2018). Implicit false belief tracking is preserved in late adulthood. *Quarterly Journal of Experimental Psychology*, 71(9), 1980-1987. Doi: 10.1177/1747021817734690

Green, M.F., Horan, W.P., & Lee, J. (2015). Social Cognition in schizophrenia. *Nat Rev Neurosci*, 16(10), 1211-1220.

Green, M.F., Horan, W. P., & Lee, J. (2019). Nonsocial and social cognition in schizophrenia: Current evidence and future directions. *World Psychiatry*, 18, 146–161. doi: 10.1002/wps.20624.

Haag, S., Haffner, P., Quinlivan, E., Brüne, M., & Stamm, T. (2016). No differences in visual theory of mind abilities between euthymic bipolar patients and healthy controls. *International Journal of Bipolar Disorders*, 4, 20.

Harvey, P. D., & Rosenthal, J. B. (2018). Cognitive and functional deficits in people with schizophrenia: Evidence for accelerated or exaggerated aging? *Schizophrenia Research*, 196, 14–21. doi: 10.1016/j.schres.2017.05.009.

Harrison, A., Sullivan, S., Tchanturia, K., & Treasure, J. (2009). Emotion Recognition and Regulation in Anorexia Nervosa. *Clinical Psychology and Psychotherapy*, 16, 348-356. Doi:10.1002/cpp.628

Hartman, L.I., Heinrichs, R.W., & Mashhadi, F. (2019). The continuing story of schizophrenia and schizoaffective disorder: One condition or two? *Schizophrenia Research: Cognition*, 16, 36-42. Doi: doi.org/10.1016/j.scog.2019.01.001

Havet-Thomassin, V., Allain, P., Etcharry-Bouyx, F., & Gall, D. Le. (2006). What about theory of mind after severe brain injury? Brain Injury, 20(1), 83-91. Doi: 10.1080/02699050500340655

Heitz, C., Noblet, V., Phillipps, C., Cretin, B., Vogt, N., Philippi, N., … Blanc, F. (2016). Cognitive and affective theory of mind in dementia with Lewy bodies and

Alzheimer's disease. *Alzheimer's Research & Therapy*, 8,10. Doi: 10.1186/s13195-016-0179-9

Helle, S., Løberg, E.M., Gjestad, R., Schnakenberg Martin, A.M., & Lysaker, P.H. (2017) The positive link between executive function and lifetime cannabis use in schizophrenia is not explained by current levels of superior social cognition. *Psychiatry Research*, 250, 92-98. Doi: dx.doi.org/10.1016/j.psychres.2017.01.047

Henry, J.D., Phillips, L., Crawford, J.R., Ietswaart, M., & Summers, F. (2006). Theory of mind following traumatic brain injury: The role of emotion recognition and executive dysfunction. *Neuropsychologia*, 44, 1623-1628.

Henry, J.D., Phillips, L, Beatty, W.W., McDonald, S., Longley, W.A., Joscelyne, A., & Rendell, P.G. (2009). Evidence for deficits in facial affect recognition and theory of mind in multiple sclerosis. Journal of the International Neuropsychological Society, 15, 277-285. Doi: doi:10.1017/S1355617709090195

Henry, J.D, Mazur, M., & Rendell, P.G. (2009). Social-cognitive difficulties in former users nof methamphetamine. *British Journal of Clinical Psychology*, 48, 323-327.

Higgins, J.P., Thompson S.G., Deeks J.J., Altman D.G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, 327, 557–60.

Hirao, K., Miyata, J., Fujiwara, J, Yamada, M., Namiki, C., Shimuzi, M., … Murai, T. (2008). Theory of mind and frontal lobe pathology in schizophrenia: a voxel-based morphometry study. *Schizophrenia Research*, 2008, 165-174.

Hoche, F., Guell, X., Sherman, J.C, Vangel, M.G., & Schmahmann, J.D. (2016). Cerebellar contribution to social cognition. Cerebellum, 15(6), 732-743. Doi: 10.1007/s12311-015-0746-9

Hotier, S., Leroy, F., Boisgontier J., Laidi, C., Mangin, J.F., Delorme, R., … Houenou, J. (2017). Social cognition in autism is associated with the neurodevelopment of

the posterior superior temporal sulcus. *Acta Psychiatrica Scandinavica*, 136, 517-525. Doi: doi.org/10.1111/acps.12814

Hysek, C.M., Domes, G., & Liechiti, M.E. (2012). MADA enhances “mind reading” of positive emotions and impairs “mind reading” of negative emotions. *Psychopharmacology*, 222, 293-302

Ignatova, V.G., Surchev, J.K., Stoyanova, T.G., Vassilev, P.M., Haralanov, L.H., & Lyudmila, T.P. (2020). Social Cognition Impairments in Patients with Multiple Sclerosis: Comparison with Grade of Disability. *Neurology India*, 68(1), 94-98. Doi: 10.4103/0028-3886.279700

Javkowiak-Siuda, K., Baron-cohen, S., Bialaszek, W., Dopierala, A., Kozlowska, A., & Pymarczsyk, K. (2016). Psychometric evaluation of the ‘Reading the Mind in the Eyes’ test with samples of different ages from a Polish population. *Studia Psychologica*, 58(1), 18-31.

Jelson-Swain, L., Persad, C., Burkard, D., & Welsh, R.C. (2015). Action Processing and Mirror Neuron Function in Patients with Amyotrophic Lateral Sclerosis: An fMRI Study. Plos One, 10(4), e0119862. Doi: 10.1371/journal.pone.0119862

Jermakow, N., & Brzezicka, A. (2016). How autistic are anorectic females? Similarities and differences between anorexia nervosa and autism spectrum disorders. *Clinical Neuropsychiatry*, 13(4/5): 53-58.

Kalbe, E., Schlegel, M., Sack, A.T., Nowak, D.A., Dafotakis, M., Bangard, C., … Kessler, J. (2010). Dissociating cognitive from affective theory of mind: a TMS study. *Cortex*, 46, 469-780. Doi: 10.1016/j.cortex.2009.07.010

Kelemen. O., Kéri. S., Must. A., Benedek. G., & Janka. Z. (2004). No evidence for impaired ‘theory of mind’ unaffected first-degree relatives of schizophrenia patients. *Acta Psychiatrica Scandinavica*, 110,146-149.

Kemmis, L., Hall, J.K., Kingston, R., & Morgan, M.J. (2007). Impaired fear recognition in regular recreational cocaine users. *Psychopharmacology*, 194, 151-159. Doi: 10.1007/s00213-007-0829-5

Kenyon, M., Samarawickrema, N., DeJong, H., Van den Eynde, F., Startup, H, Lavender, A., … Schmidt, U. (2012). Theory of Mind in Bulimia Nervosa. *International Journal of Eating Disorder*, 45(3), 377-384. Doi: doi.org/10.1002/eat.20967

Kéri, S., Kállai, I., & Csigó, K. (2020). Attribution of Mental States in Glossolalia: A Direct Comparison With Schizophrenia. *Frontiers in Psychology*, 11, 638. Doi: 10.3389/fpsyg.2020.00638

Khorashad, B.S., Khazai, B., Roshan, G.M., Hiradfar, M., Afkhamizadeh, M., & van de Grift, T.C. (2018). Prenatal testosterone and theory of mind development: Findings from disorders of sex development. *Psychoneuroendocrinology*, 89, 250-255. Doi: doi.org/10.1016/j.psyneuen.2017.12.008

Kilic, F., Demirdas, A., Isik, Ü., Akkus, M., Atay, I.M., & Kuzugüdenlioğlu, D. (2020). Empathy, Alexithymia, and Theory of Mind in Borderline Personality Disorder. *The Journal of Nervous and Mental Disease*, 208(9), 736-741. Doi: 10.1097/NMD.0000000000001196

Kim, Y.T., Kwon, D.H., & Chang, Y. (2011). Impairments of facial emotion recognition and theory of mind in methamphetamine abusers. *Psychiatric Research*, 186, 80-84. Doi: 10.1016/j.psychres.2010.06.027

Kornreich, C., Delle-Vigne, D., Knittel, J., Nerincx, A., Campanella, S., Noel, X., … Ermer, E. (2011). Impaired conditional reasoning in alcoholics: A negative impact on social interactions and risky behaviors? *Addiction*, 106(5), 951-959. Doi: 10.1111/j.1360-0443.2010.03346.x

Köther, U., Veckenstedt, R., Vitzthum, F., Roesch-Ely, D., Pfueller, U., Scheu, F., & Moritz, S. (2012). "Don't give me that look"—Overconfidence in false mental state perception in schizophrenia. *Psychiatry Research*, 196, 1-8. Doi: 10.1016/j.psychres.2012.03.004

Lam, B.Y., Raine, A., & Lee, T.M. (2014). The relationship between neurocognition and symptomatology in people with schizophrenia: social cognition as the mediator. *BMC Psychiatry*, 14, 138.

Launay, J., Pearce, E., Wlodarski, R., van Duijn, M., Carney, J., & Dunbar, R.I.M. (2015). Higher-order mentalizing and executive functioning. *Personality and Individual Differences*, 86, 6-14. Doi: dx.doi.org/10.1016/j.paid.2015.05.021

Lawson, R. (2015). I just love the attention: implicit preference for direct eye contact. Visual Cognition, 23(4), 450-488. Doi: 10.1080/13506285.2015.1039101

Lee, H.S., Corbera, S., Poltorak, A., Park, K., Assaf, M., Bell, M.D., … Choi, K.H. (2018). Measuring theory of mind in schizophrenia research: Cross-cultural validation. *Schizophrenia Research*, 201, 187-195. Doi: doi.org/10.1016/j.schres.2018.06.022

Lee, H.R., Nam, G, & Hur, J.W. (2020). Development and validation of the Korean version of the Reading the Mind in the Eyes Test. *Plos One*, 15(8), e0238309. Doi: 10.1371/journal.pone.0238309

Lee, S.A., Guajardo, N.R., Short, S.D., & King, W. (2010). Individual difference in ocular level empathic accuracy ability: the predictive power of fantasy empathy. *Personality and Individual Differences*, 49, 68-71.

Leppanen, L., Cardi, V., Ng, K.W., Paloyelis, Y., Stein, D., Tchanturia, K., & Treasure, J. (2017). Effects of Intranasal Oxytocin on the Interpretation and Expression of Emotions in Anorexia Nervosa. *Journal of Neuroendocrinology*, 29, 1-13. Doi: 10.1111/jne.12458

Lopez-del-Hoyo, Y., Panzano, M.G., Lahera, G., Herrera-Mercadal, P., Navarro-Gil, M., Campos, D. Borao, L., … García-Campayo, J. (2019). Difference between individuals with schizophrenia or obsessive-compulsive disorder and healthy controls in social cognition and mindfulness skills: a controlled study. *Plos One*. 14(12), e0225608. Doi: doi.org/10.1371/journal.pone.0225608

López-Navarro, E. (2018). Contributions of executive functions components to affective and cognitive theory of mind in outpatients diagnosed with schizophrenia. *Psychiatry Research*, 269, 124-125. Doi: doi.org/10.1016/j.psychres.2018.08.018

Ma, J., Guo, Q.H., Zhang, Y.Y., & He, M.Q. (2015). Effect of Chinese version of "Reading the Mind in the Eyes" test on the diagnosis of mild cognitive impairment and Alzheimer's dementia. *Chin J Geristr*, 34(11), 1218-1222. Doi: 10.3760/cma.j.issn.0254-9026.2015.11.018 [in Chinese]

Martinez-Sanchez, F., Fernández-Abscal, E.G., & Sánchez-Pérez, N. (2017). Recognition of Emotional Facial Expression in Alexithymia. *Studia Psychologica*, 59(3), 206-216. Doi: 10.21909/sp.2017.03.741

McPhilemy, G., Nabulsi, L., Kilmartin, L., Whittaker, J., Martyn, F.M., Hallahan, B., … Cannon, D.M. (2020). Resting-State Network Patterns Underlying Cognitive Function in Bipolar Disorder: A Graph Theoretical Analysis. *Brain Connectivity,* 10(7), 355-367. Doi: doi.org/10.1089/brain.2019.0709

Medina-Pradas, C., Navarro, J.B., Álvarez-Moya, E.M., Grau, A., & Obiols, J.E. (2012). Emotional theory of mind in eating disorders. *International Journal of Clinical and Health Psychology*, 12(2), 189-202.

Melchers, M., Montag, C., Markett, S., & Reuter, M. (2015). Assessment of empathy via self-report and behavioural paradigms: data on convergent and discriminant validity. *Cognitive Neuropsychiatry*, 20(2), 157-171. Doi: dx.doi.org/10.1080/13546805.2014.991781

Melchers, M., Montag, C., Reuter, M., Spinath, F.M., & Hahn, E. (2016). How heritable is empathy? Differential effects of measurement and subcomponents. *Motiv Emot*, 40, 720-730. Doi: 10.1007/s11031-016-9573-7.

Mesholam-Gately, R.I., Giuliano, A.J., Goff, K.P., Faraone, S.V., & Seidman, L.J. (2009). Neurocognition in first-episode schizophrenia: A meta-analytic review. *Neuropsychology,* 23(3), 315–336. Doi: 10.1037/a0014708.

Meyer, J., & Shean, G. (2006). Social-Cognitive Functioning and Schizotypal Characteristics. *The Journal of Psychology*, 140(3), 199-207.

Moieni, M., Irwin, M.R., Jevtic, I., Breen, E.C., & Eisenberger, N.I. (2015) Inflammation impairs social cognitive processing: a randomized controlled trial of endotoxin. *Brain Behav Immum*, 48, 132-138.

Montag, C., Haase, L., Seidel, D., Bayerl, M., Gallinat, J., Herrmann, U., & Dannecker, K. (2014). A pilot RCT of Psychodynamic Group Art Therapy for Patients in Acute Psychotic Episodes: Feasibility, Impact on Symptoms and Mentalising Capacity. *Plos One*, 9(11), e112348. Doi: 10.1371/journal.pone.0112348

Murphy, D. (2006). Theory of mind in Asperger's syndrome, schizophrenia and personality disordered forensic patients. *Cognitive Neuropsychiatry*, 11(2), 99-111. Doi: https://doi.org/10.1080/13546800444000182

Murphy, D. (2007). Theory of mind functioning in mentally disordered offenders detained in high security psychiatric care: its relationship to clinical outcome, need and risk. *Criminal Behaviour and Mental Health*, 17, 300-311. Doi: 10.1002/cbm.664

Nandrino, J.L., Gandolphe, M.C., Alexandre, C., Kmiecik, E., Yguel, J., & Urso, L. (2014). Cognitive and affective theory of mind abilities in alcohol-dependent patients: The role of autobiographical memory. Drug and Alcohol Dependence, 143, 65-73. Doi: dx.doi.org/10.1016/j.drugalcdep.2014.07.010

Navarra-Ventura, G., Fernandez-Gonzalo, S., Turon, M., Pousa, E., Palao, D., Cardoner, N., & Jodar, M. (2018). Gender Differences in Social Cognition: A Cross-Sectional Pilot Study of Recently Diagnosed Patients with Schizophrenia and Healthy Subjects. *The Canadian Journal of Psychiatry*, 63(8), 538-546. Doi: 10.1177/0706743717746661

Nejati, V., Zabihzadeh, A., Maleki, G., & Tehranchi, A. (2012). Mind reading and mindfulness deficits in patients with major depression disorder. *Procedia—Social and Behavioral Sciences*, 32, 431-437. Doi: 10.1016/j.sbspro.2012.01.065

Newbury-Helps, J., Feigenbaum, J., & Fonagy, P. (2017). Offenders with Antisocial Personality Disorder Display More Impairments in Mentalizing. *Journal of Personality Disorders*, 31(2), 232-255.

Nietlisbach, G., Maercker, A., Rössler, W., & Haker, H. (2010). Are empathic abilities impaired in posttraumatic stress disorder? Psychological Reports, 106(3), 832-844. Doi: 10.2466/PR0.106.3.832-844

Okruszek, T., Bala, A., Wordecha, M., Jarkiewicz, M., Wysokinski, A., Szczepocka, E., … Marchel, A. (2016). Social cognition in neuropsychiatric populations: a comparison of theory of mind in schizophrenia and mesial temporal lobe epilepsy. *Scientific Reports*, 7: 484. Doi: 10.1038/s41598-017-00565-2

Oldershaw, A., Hambbook, D., Tchanturia, K., Treasure, J., & Schmidt, U. (2010). Emotional Theory of Mind and Emotional Awareness in Recovered Anorexia Nervosa Patients. *Psychosomatic Medicine*, 72, 73-79. Doi: 0033-3174/10/7201-0073

Pahnke, R., Mau-Moeller, A., Hamm, A.O., & Lischke, A. (2020). Reading the Mind in the Eyes of Children Test (RME-C-T): Development and Validation of a Complex Emotion Recognition Test. *Frontiers in Psychiatry*, 11, 376. Doi: 10.3389/fpsyt.2020.00376

Pardini, M., & Nichelli, P.F. (2009). Age-Related Decline in Mentalizing Skills Across Adult Life Span. *Experimental Aging Research*, 35(1), 98-106. Doi: 10.1080/03610730802545259

Pardini, M., Gialloreti, L.E., Mascolo, M., Benassi, F., Abate, L, Guida, S., … Cocito, L. (2013) Isolated theory of mind deficits and risk for frontotemporal dementia: a longitudinal pilot study. J *Neurol Neurosurg Psychiatry*, 84, 818-821. Doi: dx.doi.org/10.1136/jnnp-2012-303684

Pentaraki, A.D., Stefanis, N.C., Stahl, D., Theleritis, C., & Toulopoulou, T., (2012). Theory of Mind as a potential trait marker of schizophrenia: A family study. *Cognitive Neuropsychiatry*, 17(1), 64-89.

Petersen, R., Brakoulia, V., & Langdon, R. (2016). An experimental investigation of mentalization ability in borderline personality disorder. *Comprehensive Psychiatry*, 64, 12-21. Doi: dx.doi.org/10.1016/j.comppsych.2015.10.004

Petroni, A., Canales-Johnson, A., Urquina, H., Guex, R., Hurtado, E., Blenkmann, A., … Ibañez, A. (2011). The cortical processing of facial emotional expression is associated with social cognition skills and executive functioning: A preliminary study. *Neuroscience Letters*, 505, 41-46. Doi: 10.1016/j.neulet.2011.09.062

Philippe, A., Havet-Thomassin, V., Verny, C., Gohier, B., Lancelot, C., Besnard, J., … Le Gall, D. (2011). Evidence for Deficits on Different Components of Theory of Mind in Huntington's Disease. *Neuropsychology*, 25(6), 741-751. Doi: 10.1037/a0024408

Pinkham, A.E., Harvey, P.D., & Penn, D.L. (2018). Social Cognition Psychometric Evaluation: Results of the Final Validation Study. *Schizophrenia Bulletin*, 44(4), 737-748. Doi: 10.1093/schbul/sbx117

Pino, M.C., De Berardis, D., Mariano, M., Vellante, F., Serroni, N., Valchera, A., … Mazza, M. (2016). Two systems for empathy in obsessive-compulsive disorder: mentalizing and experience sharing. *Revista Brasileira de Psiquiatria*, 38, 307-313. Doi: 10.1590/1516-4446-2015-1679

Poletti, M., & Bonuccelli, U. (2013). Alteration of affective Theory of Mind in amnestic mild cognitive impairment. *Journal of Neuropsychology*, 7, 121-131. Doi: 10.1111/j.1748-6653.2012.02040.x

Poletti, M, Vergallo, A., Ulivi, M, Sonnoli, A., & Bonuccelli, U. (2013). Affective theory of mind in patients with Parkinson's disease. *Psychiatry and Clinical Neurosciences*, 67, 273-273.

Preller, K.H, Hulka, L.M, Yonmoos, M., Jenni, D., Baumgartner, M.R., Seifritz, E., … Quednow, B.B. (2013). Impaired emotional empathy and related social network deficits in cocaine users. *Addiction Biology*, 19: 452-466. Doi: 10.1111/adb.12070

Preti, A., Vellante, M., & Petretto, D.R. (2017). The psychometric properties of the "Reading the Mind in the Eyes" Test: an item response theory (IRT) analysis. *Cognitive Neuropsychiatry*, 22(3), 233-253. Doi: dx.doi.org/10.1080/13546805.2017.1300091

Prevost, M., Carrier, M.E., Chowne, G., Zelkowitz, P., Joseph, L., & Gold, I. (2014). The Reading the Mind in the Eyes test: validation of a French version and exploration of cultural variations in a multi-ethnic city. *Cognitive Neuropsychiatry*, 19(3), 189-204. Doi: dx.doi.org/10.1080/13546805.2013.823859

Purcell, A.L., Phillips, M., & Gruber, J. (2013). In Your Eyes: Does Theory of Mind Predict Impaired Life Functioning in Bipolar Disorder? *J Affect Disord*, 151(3), 1-15. Doi: 10.1016/j.jad.2013.06.051

Quintana, D.S., Guastella, A.J., Outhred, T., Hickie, I.B., & Kemp, A.H. (2012). Heart rate variability is associated with emotion recognition: Direct evidence for a relationship between the autonomic nervous system and social cognition. *International Journal of Psychophysiology*, 86, 168-172.

Radke, S., & de Bruijn, E.R.A. (2015). Does oxytocin affect mind-reading? A replication study. *Psychoneuroendocrinology*, 60, 75-81. Doi: dx.doi.org/10.1016/j.psyneuen.2015.06.006

Raimo, S., Trojano, L., Pappacena, S., Alaia, R., Spitaleri, D., Grossi, D., & Santangelo, G. (2017). Neuropsychological Correlates of Theory of Mind Deficits in Patients With Multiple Sclerosis. *Neuropsychology*, 31(7), 811-821. Doi: dx.doi.org/10.1037/neu0000372

Redondo, I., & Herrero-Fernández, D. (2018). Validation of the Reading the Mind in the Eyes Test in a healthy Spanish sample and women with anorexia nervosa. *Cognitive Neuropsychiatry*, 23(4), 201-217. Doi: doi.org/10.1080/13546805.2018.1461618

Reynolds, M.T, Van Rheenen, T.E., & Rossell, S.L. (2014). Theory of mind in first degree relatives of individuals with bipolar disorder. *Psychiatry Research*, 219, 400-402. Doi: dx.doi.org/10.1016/j.psychres.2014.05.041

Robinson, L.J., Gary, J.M., Burt, M., Ferrier, I.N., & Gallagher, P. (2015). Processing of Facial Emotion in Bipolar Depression and Euthymia. Journal of the International Neuropsychological Society, 21, 709-721. Doi: doi.org/10.1017/S1355617715000909

Rodrigues, S.M., Saslow, L.R., Gracia, N., John, O.P, & Keltner, D. (2009). Oxytocin receptor genetic variation relates to empathy and stress reactivity in humans. PNAS, 106(50), 21437-21441. Doi: www.pnas.org/cgi/doi/10.1073/pnas.0909579106

Rominger, C., Bleier, A., Fitz, W., Marksteiner, J., Fink, A., Papousek, I., & Weiss, E.M. (2016). Auditory top-down control and affective theory of mind in schizophrenia with and without hallucinations. Schizophrenia Research, 174(1-3), 192-196. doi: dx.doi.org/10.1016/j.schres.2016.05.006

Romosan, A.M., Dehelean, L., Romosan, R.S., Andor, M., Bredicean, A.C., & Simu, M.A. (2019). Affective theory of mind in Parkinson's disease: the effect of cognitive performance. *Neuropsychiatric Disease and Treatment*, 15, 2521-2535. Doi: 10.2147/NDT.S219288

Russel, T.A., Schmidt, U., Doherty, L., Young, V., & Tchanturia, K. (2009). Aspects of social cognition in anorexia nervosa: affective and cognitive theory of mind. *Psychiatry Research*, 168, 181-185. Doi: 10.1016/j.psychres.2008.10.028

Sacchetti, S., Robinson, P., Bogaardt, A., Clare, A., Ouellet-Courtois, C., Luyten, P., … Fonagy, P. (2019). Reduced mentalizing in patients with bulimia nervosa and features of borderline personality disorder: A case-control study. *BMC Psychiatry*, 19, 134. Doi: doi.org/10.1186/s12888-019-2112-9

Sanvicente-vieira, B., Kluwe-schiavon, B., Corcoran, R., & Grassi-oliveira, R., (2017). Theory of Mind Impairments in Women with Cocaine Addiction. *J Stud Alcohol Drug*, 78(2), 258-267. Doi: 10.15288/jsad.2017.78.258

Sabater, A., Gracía-Blanco, A.C., Verdet, H.M., Sierra, P., Ribes, J., Villar, I., … Livianos, L. (2016). Comparative neurocognitive effects of lithium and anticonvulsants in long-term stable bipolar patients. *Journal of Affective Disorders*, 190, 34-40. Doi: dx.doi.org/10.1016/j.jad.2015.10.008

Scherzer, P., Achim, A., Leveille, E., Boisseau E., & Stip E. (2015). Evidence from paranoid schizophrenia for more than one component of theory of mind. *Frontiers in Psychology*, 6, 1643. Doi: 10.3389/fpsyg.2015.01643.

Schilling, L., Wingenfeld, K., Löwe, B., Moritz, S., Terfehr, K., Köther, U., & Spitzer, C. (2012). Normal mind-reading capacity but higher response confidence in borderline personality disorder patients. *Psychiatry and Clinical Neurosciences*, 66, 322-327. Doi: 10.1111/j.1440-1819.2012.02334.x

Schimit, J.Z., & Zachariae, R. (2009). PTSD and Impaired Eye Expression Recognition: A Preliminary Study. *Journal of Loss and Trauma*, 14(1), 46-56. Doi: http://dx.doi.org/10.1080/15325020802537096

Schneider, A., Johnston, C., Tassone, F., Sansone, S., Hagerman, R.J., Ferrer, E., … Hessl, D. (2016). Broad Autism Spectrum and Obsessive Complusive Symptoms in Adults with the Fragile X Premutation. *Clin Neuropsychol*, 30(6), 929-943. Doi: 10.1080/13854046.2016.1189536

Schuwerk, T., Vuori, M., & Sodian, B. (2015). Implicit and explicit Theory of Mind reasoning in autism spectrum disorders: The impact of experience. *Austim*, 19(4), 459-468. Doi: 10.1177/1362361314526004

Sheffield, J. M., Karcher, N. R., & Barch, D. M. (2018). Cognitive deficits in psychotic disorders: A lifespan perspective. *Neuropsychology Review*, 28, 509–533. Doi: 10.1007/s11065-018-9388-2.

Simon, M., Németh, N., Gálber, M., Lakner, E., Csernela, E., Tényi, T., & Czéh, B. (2019). Childhood Adversity Impairs Theory of Mind Abilities in Adult Patients with Major Depressive Disorder. *Frontiers in Psychiatry*, 10, 867. Doi: 10.3389/fpsyt.2019.00867

Slessor, G., Phillips, L.H, & Bull, R. (2007). Exploring the Specificity of Age-Related Differences in Theory of Mind Tasks. *Psychology and Aging*, 22(3), 639-643. Doi: 10.1037/0882-7974.22.3.639

Smeets, T. Dziobek, I., & Wolf, O.T. (2009). Social cognition under stress: Differential effects of effects of stress-induced cortisol elevations in healthy young men and women. *Hormones and Behavior*, 55(4), 507-513. Doi: 10.1016/j.yhbeh.2009.01.011

Standford, A.D., Messinger, J., Malaspina, D., & Corcoran, C.M. (2011). Theory of Mind in Patients at Clinical High Risk for Psychosis. *Schizophrenia Research*, 131(1-3), 11-17. Doi: 10.1016/j.schres.2011.06.005

Sun, F., Lin, L.Z., Dai, M.X., Luo, M.F, Guo, C.H., Yang, H.W., & Jing, J. (2017). Autistic Traits, Empathy and Theory of Mind of College Students with Different Level

of Prosocial Behaviors. Journal of Sun Yat-Sen University (Medical Sciences), 38(6): 938-960. [in Chinese]

Szanto, K., Dombrovski, A.Y., Sahakian, B.J., Mulsant, B.H., Houck, P.R., Reynolds, C.F., & Clark, L. (2012). Social emotion recognition, social functioning, and attempted suicide in late-life depression. *Am J Geriatr Psychiatry*, 20(3), 257-265. Doi: 10.1097/JGP.0b013e31820eea0c.

Tapajóz, F., Soneira, S., Catoira, N., Aulicino, A., & Allegri, R.F. (2019). Impaired theory of mind in unaffected first-degree relatives of patients with anorexia nervosa. *Eur Eat Disorders Rev*, 27: 692-699.

Tella, M.D., Castelli, L., Colonna, F., Fusaro, E., Torta, R., Ardito, R.B, & Adenzato, M. (2015). Theory of Mind and Emotional Functioning in Fibromyalgia Syndrome: An Investigation of the Relationship between Social Cognition and Executive Function. *Plos One*, 10(1), e0116542. Doi: 10.1371/journal.pone.0116542

Thaler, N.S., Allen, D.N., Sutton, G.P., Vertinski, M., & Ringdahl, E.N. (2013). Differential impairment of social cognition factors in bipolar disorder with and without psychotic features and schizophrenia. *Journal of Psychiatric Research*. Doi: http://dx.doi.org/10.1016/j.jpsychires.2013.09.010

Thoma, P., Winter, N., Juckel, G., & Roser, P. (2013). Mental State decoding and mental state reasoning in recently detoxified alcohol-dependent individuals. *Psychiatry Research*, 205, 232-240. Doi: http://dx.doi.org/10.1016/j.psychres.2012.08.042

Tylec, A., Jeleniewicz, W., Mortimer, A., Bednarska-Makaruk, M., & Kucharska, K. (2017). Interaction Between Val158Met Catechol-O-Methyltransferase Polymorphism and Social Cognitive Functioning in Schizophrenia: Pilot Study. *Annals of Human Genetics*, 81, 267-275. Doi: 10.1111/ahg.122209

Tso, I.F., Grove, T.B., & Taylor, S.F. (2010). Emotional experience predicts social adjustment independent of neurocognition and social cognition in schizophrenia. *Schizophrenia Research*, 122, 156-163.

Uhlmann, A., Ipser, J.C., Wilson, D., & Stein, D.J. (2018). Social cognition and aggression in methamphetamine dependence with and without a history of psychosis. *Metabolic Brain Disease*, 33, 559-568. Doi: doi.org/10.1007/s11011-017-0157-3

Ünal-Aydin, P., Balıçı, K., Sönmez, I. & Aydin, O. (2020). Associations between emotion recognition and social networking site addiction. *Psychiatry Research*, 284, 112673. Doi: 10.1016/j.psychres.2019.112673

Viechtbauer, W. (2005). Bias and Efficiency of Meta-Analytic Variance Estimators in the Random-Effects Model. *Journal of Educational and Behavioral Statistics*, 30(3), 261-293.

Villa, J.M, Ganzel, B.L, Yoder, K.J., Chen, G.M., Lyman, L.T., Sidari, A.P., … Belmonte, M.K. (2010). More than maths and mindreading: sex differences in empathizing/systemizing covariance. *Autism Research*, 3, 174-184. Doi: doi.org/10.1002/aur.143

von Elm, E., Altman, D.G., Egger, E., Pocock, S.J., Gøtzsche, P.C. Vandenbrouche, J.P., & STROBE Initiative. (2008). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Chin Epidemiol, 61(4), 344-349. Doi: 10.1016/j.jclinepi.2007.11.008

Wang, X., Song, Y.Y., Zhen, Z.L., & Liu, J. (2016). Functional Integration of the Posterior Superior Temporal Sulcus Correlates with Facial Expression Recognition. *Human Brain Mapping*, 37, 1930-1940. Doi: 10.1002/hbm.23145

Wang, Y.R. (2018). tDCS for cognitive deficits in patients with schizophrenia: a randomized controlled trial. [master dissertation, Shanghai Jiaotong University] [in Chinese]

Weisman, O., Pelphrey, K.A., Leckman, J.F., Feldman, R., Lu, Y.F., Chong, A., … Ebstein, R.P. (2015). The association between 2D: 4D ratio and cognitive empathy is contingent on a common polymorphism in the oxytocin receptor gene (OXTR rs53576). *Psychoneuroendocrinology*, 58, 23-32. Doi: dx.doi.org/10.1016/j.psyneuen.2015.04.007

Whitton, A.E., & Henry, J.D. (2013). The relationship between sub-clinical obsessive-compulsive symptoms and social cognition in chronic schizophrenia. *British Journal of Clinical Psychology*, 52(2), 115-128. Doi: 10.1111/bjc.12004

Wilu Wilu, A., Allain, P., Moustafa, A.A., & El Haj, M. (2019). "To whom did I tell that information?": Relationship between destination memory and theory of mind in traumatic brain injury. *Applied Neuropsychology Adult*, 28(1), 1-11. Doi: 10.1080/23279095.2019.1585350

Wolkenstein, L., Schönenberg, M., Schirm, E., & Hautzinger, M. (2011). I can see what you feel, but I can't deal with it: Impaired theory of mind in depression. *Journal of Affective Disorders*, 132, 104-111. Doi: 10.1016/j.jad.2011.02.010

Woolley, J.D., Chuang, B., Lam, O., Lai, W., O'Donovan, A., Rankin, K.P., … Vinogradov, S. (2014). Oxytocin administration enhances controlled social cognition in patients with schizophrenia. *Psychoneuroendocrinology*, 47, 116-125. Doi: 10.1016/j.psyneuen.2014.04.024

Yang, C.C., Khalifa, N., Lankappa, S., & Völlm, B. (2018). Effects of intermittent theta burst stimulation applied to the left dorsolateral prefrontal cortex on empathy and impulsivity in healthy adult males. *Brain and Cognition*, 128, 37-45. Doi: doi.org/10.1016/j.bandc.2018.11.003

Zabihzadeh, A., Maleki, G., Richman, M.J., Hatami, A., Alimardani, Z., & Heidari, M. (2017). Affective and cognitive theory of mind in borderline personality disorder: The role of comorbid depression. *Psychiatry Research*, 257, 144-149. Doi: dx.doi.org/10.1016/j.psychres.2017.07.034

Zegarra-Valdivia, J.A., & Vilca, B.N.C. (2019). Social cognition and executive function in borderline personality disorder: evidence of altered cognitive process. *Salud Mental*, 42(1), 33-41. Doi: 10.17711/SM.0185-3325.2019.005

Zhang, W. (2010). A Study on the Characteristics of College Students' Theory of Mind and Its Relationship with Altruistic Tendency. [Master dissertation, Henan University] [in Chinese]

Zilioli, S., Ponzi, D., Henry, A., & Maestripieri, D. (2015). Testosterone, Cortisol and Empathy: Evidence for the Dual-Hormone Hypothesis. *Adaptive Human Behavior and Physiology*, 1, 421-433. Doi: 10.1007/s40750-014-0017-x

3.7 Supplementary materials of chapter three

APPENDIX ONE: Search strategies in different databases

Pubmed: "RMET"[All Fields] OR (("reading"[MeSH Terms] OR "reading"[Title/Abstract] OR "readings"[Title/Abstract]) AND ("sci am mind"[Journal] OR "mind"[Journal] OR "mind"[All Fields]) AND ("eye"[MeSH Terms] OR "eye"[Title/Abstract] OR "eyes"[Title/Abstract])) OR ((((("schizophrenia"[MeSH Terms] OR "schizophrenia"[Title/Abstract] OR "schizophrenias"[Title/Abstract] OR "schizophrenia s"[Title/Abstract]) AND (("eye"[MeSH Terms] OR "eye"[Title/Abstract]) AND ("research design"[MeSH Terms] OR ("research"[Title/Abstract] AND "design"[Title/Abstract]) OR "research design"[Title/Abstract] OR "test"[Title/Abstract]))) NOT "RMET"[Title/Abstract]) NOT (("reading"[MeSH Terms] OR "reading"[Title/Abstract] OR "readings"[Title/Abstract]) AND ("sci am mind"[Journal] OR "mind"[Journal] OR "mind"[Title/Abstract]) AND ("eye"[MeSH Terms] OR "eye"[Title/Abstract] OR "eyes"[Title/Abstract])))
Search date: 6th Aug. 2020

Web of science: (AB=RMET) OR (AB =reading the mind in the eyes) OR (AB=reading the mind in the eye)OR ((AB =schizophrenia) AND (AB =eye test) NOT ((AB =RMET) OR (AB =reading the mind in the eyes) OR (AB=reading the mind in the eye)))
Search date: 6th Aug. 2020

Psycinfo/EBSCO: AB RMET OR AB reading the mind in the eyes OR AB reading the mind in the eye OR (AB schizophrenia) AND (AB eye test) NOT ((AB RMET) OR (AB reading the mind in the eyes) OR (AB reading the mind in the eye)))
Search date: 6th Aug. 2020

CNKI: TKA=RMET OR TKA=读眼识心 OR TKA=读眼阅读 OR TKA=眼区读心 OR TKA=眼神阅读 OR TKA=Reading the Mind in the Eyes OR TKA=reading the Mind in the Eye OR (TKA=schizophrenia AND (TKA=眼区 OR TKA=眼神) NOT TKA=RMET NOT TKA=reading the mind in the eyes NOT TKA=reading the mind in the eye)
Search date: 11th Aug. 2020

Wanfang: 摘要:(RMET) or 摘要:('读眼识心') or 摘要:('读眼阅读') or 摘要:('眼区读心') or 摘要:('眼神阅读') or 摘要:('Reading the Mind in the Eyes') or 摘要:('reading the Mind in the Eye') or (摘要:(schizophrenia) and (摘要:('眼区') or 摘要:('眼神')) not 摘要:(RMET) not 摘要:('reading the mind in the eyes') not 摘要:('reading the mind in the eye'))
Search date: 11th Aug. 2020

Chapter Four: Comparison of social cognition using an adapted Chinese version of the Reading the Mind in the Eyes Test in drug-naive and regularly medicated individuals with chronic schizophrenia and healthy controls in rural China

Abstract

Background. Social cognition has not previously been assessed in treatment-naive patients with chronic schizophrenia, in patients over 60 years of age, or in patients with less than 5 years of schooling.

Methods. We revised a commonly used measure of social cognition, the Reading the Mind in the Eyes Test (RMET), by expanding the instructions, using both self-completion and interviewer-completion versions (for illiterate respondents), and classifying each test administration as 'successfully completed' or 'incomplete'. The revised instrument (RMET-CV-R) was administered to 233 treatment-naive patients with chronic schizophrenia (UT), 154 treated controls with chronic schizophrenia (TC), and 259 healthy controls (HC) from rural communities in China.

Results. In bivariate and multivariate analyses, successful completion rates and RMET-CV-R scores (percent correct judgments about emotion exhibited in 70 presented slides) were highest in HC, intermediate in TC, and lowest in UT (adjusted completion rates, 95.7%, 69.5%, and 50.4%, respectively; adjusted RMET-CV-R scores using GLS model, 47.0%, 40.1%, and 36.3%, respectively; adjusted RMET-CV-R scores using Heckman estimation including all the participants could not complete RMET-CV-R, 44.5%, 31.3%, and 23.1%, respectively; all $p<0.03$). Stratified analyses by method of administration (self-completed v. interviewer-completed) and by education and age ('educated-younger' v. 'undereducated-older') show the same relationship

between groups (i.e., NC>TC>UT), though not all differences remain statistically significant.

Conclusions. We find poorer social cognition in treatment-naive than in treated patients with chronic schizophrenia. The discriminant validity of RMET-CV-R in undereducated, older patients demonstrates the feasibility of administering revised versions of RMET to patients who may otherwise be considered ineligible due to education or age by changing the method of test administration and carefully assessing respondents' ability to complete the task successfully.

Keywords: Reading Mind in the Eyes Test (RMET), social cognition, schizophrenia, China, low- and middle-income countries (LMIC)

4.1 Introduction

Individuals with psychotic disorders such as schizophrenia typically exhibit neurocognitive deficits in overall cognitive abilities, attention, working memory, declarative memory, and executive functioning (Green, Horan, & Lee, 2019; Harvey, & Rosenthal, 2018; Sheffield, Karcher, & Barch, 2018; Mesholam-Gately, Giuliano, Goff, Faraone, & Seidman, 2009). Frequently, these individuals also demonstrate deficits in 'social cognition', which refers to a multidimensional set of cognitive abilities that are necessary to learn social norms and to perceive emotions and other social cues in interpersonal interactions. Importantly, impairments in social cognitive domains may impair social functioning more than other types of neurocognitive deficits (Fett, Viechtbauer, Dominguez, Penn, van Os, & Krabbendam, 2011; Green, Horan, & Lee, 2015; Green et al., 2019; Harvey, Strassnig, & Silberstein, 2019).

Current research on social cognition in schizophrenia is primarily conducted in high-income countries with urban samples comprised of young adults with good education who are currently in treatment (Bell, Corbera, Johannesen, Fiszhon, & Wexler, 2013; Bora, Yucel, & Pantelis, 2009). Much less is known about social cognition among individuals with schizophrenia in low- and middle-income countries (LMIC), particularly those living in rural communities where the level of education is low, where many community residents with psychosis are older adults, and where a substantial proportion of individuals with psychosis have never received antipsychotic medication. Another limitation of current research is that many studies use instruments that assess social cognition by asking respondents to make comments on stories or case vignettes (e.g., Mayer-Salovey-Caruso Emotional Intelligence Test MSCEIT test [Mayer, Salovey, Caruso, & Sitarenios, 2003], Hinting Task test [Corcoran, Mercer, & Frith, 1995] and the Combined Story Test [Achim, Ouellet, Roy, & Jackson, 2012]), some of which make cultural assumptions or describe situations that are unfamiliar to rural respondents from LMIC.

Another commonly used test of social cognition, the Reading the Mind in the Eyes Test (RMET) (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001), is not

based on interpreting stories or vignettes, so it is potentially less culture-dependent and, thus, may be more suitable for use in LMIC. Based on the first stage of the Theory of Mind, RMET assesses an individual's use of social cues to attribute another person's mental state. Respondents are shown images of the eye region of 36 Caucasian faces along with 4 adjectives describing contrasting mental states (e.g., terrified, upset, arrogant, annoyed) and asked to select the label that best reflects the mental state of each individual pictured and (to ensure the respondent is attending to the task) to indicate the individual's gender. The psychometric properties of RMET are good (Pinkham, Penn, Green, & Harvey, 2016; Eddy, 2019), but some authors report that it has poor construct validity because – like many other cognitive measures – it assesses a multifactorial construct (Carter, Barch, Gur, Gur, Pinkham, & Ochsner, 2009; Olderbak, Wilhelm, Olaru, Geiger, Brenneman, & Roberts, 2015). The RMET has been translated into 21 languages.

Although originally developed for Autism Spectrum Disorders, RMET has also been widely used to assess social cognition in schizophrenia (Balogh, Égerházi, Berecz, & Csukly, 2014; Caletti et al., 2013; Couture, Penn, Addington, Woods, & Perkins, 2008; Donohoe et al., 2012; Guastella et al., 2013; Kettle, O'Brien-Simpson, & Allen, 2008; Mazza et al., 2013; Stanford, Messinger, Malaspina, & Corcoran, 2011; Thaler, Allen, Sutton, Vertinski, & Ringdahl, 2013; Vohs et al., 2014; Zhang et al., 2016). However, the small sample sizes in these studies (only one had more than 100 subjects) make it impractical to adjust RMET results for important potential confounders such as age, gender, and years of schooling; only one of the studies made such adjustments (Thaler et al., 2013). Moreover, these studies were all conducted in urban areas among young or middle-aged adults (range in mean age: 23-48) with relatively high levels of education (at least middle school) who were regularly receiving antipsychotic medications. The social cognition of patients with schizophrenia who are older, have little or no formal schooling, live in rural communities, or have never been treated with antipsychotic medication have not, to our knowledge, previously been assessed.

Using a parallel Chinese version of RMET with 34 pictures of eyes of Asian individuals, Wang and colleagues reported that social cognition was worse in individuals with schizophrenia than in healthy controls (Wang, Wang, Zhu, & Wang, 2006). Another research team in China combined a translation of the 36-item English version with the 34-item Chinese version (Zhang et al., 2016; Zhang et al., 2018) into a combined 70-item Chinese version of RMET (RMET-CV). Using this instrument with urban, well-educated respondents, they found that compared to healthy controls, social cognitive deficits are greater in patients with chronic schizophrenia and less severe in individuals at clinical high risk (CHR) for schizophrenia.

This paper reports on the administration of the 70-item RMET-CV in 234 untreated patients with chronic schizophrenia (UT), 154 treated patients with chronic schizophrenia (TC), and 259 healthy controls (HC), most of whom reside in rural communities in China. Included participants are quite distinct from typical samples used in previous studies: they had a mean (sd) age of 51.1 (11.7) years and a mean of 4.0 (3.6) years of schooling (30% never attended school). Revisions in the administration and scoring of the RMET-CV (without changing the test items themselves) were required to make it suitable for use with these subjects; we call our revised scale the 'revised Chinese version of RMET' (RMET-CV-R).

In addition to confirming previous results about deficits in social cognition among patients with schizophrenia in this unique sample of respondents, this analysis aims to assess several previously unexplored issues: 1) differences in social cognition between never-treated and currently treated patients, 2) the feasibility of administering RMET-CV-R as an interviewer-completed instrument, and 3) the feasibility of administering RMET-CV-R to older respondents and respondents with little or no formal education—respondents not included in previous studies with RMET or RMET-CV.

4.2 Methods

The administration of the RMET-CV-R reported in this paper was part of an ongoing study comparing cognitive functioning among UT, TC, and HC. All participants were community residents primarily living in rural parts of the Ningxia Hui Autonomous Region (a province in northwest China) and the Guangxi Zhuang Autonomous Region (a province in southwest China). Cognitive function was assessed using an adapted Chinese version of the Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) Consensus Cognition Battery (MCCB-CV-R; Stone et al., 2020). Pilot testing revealed that rural residents with limited education were unfamiliar with the situations described in the MSCEIT vignettes used to assess social cognition in the original Chinese version of MCCB (Shi et al., 2015), so MSCEIT was replaced with RMET-CV-R in the MCCB-CV-R.

4.2.1 Participants

UT and TC were identified from the provincial registries of individuals with severe mental disorders in Ningxia and Guangxi, which are part of the National Continuing Management and Intervention Program for Psychoses (Wu et al., 2019), which aims to register and treat all persons with severe mental disorders throughout China. HC were identified from individuals registered in local health clinics (which register all permanent residents) in Ningxia.

Preliminary screening for eligibility and willingness to participate was checked by phone for all three groups (UT, TC, HC), followed by a home or local clinic visit to confirm study eligibility. During the face-to-face visit and after receiving written consent from the participant (and an adult family member for UT and TC), the adapted Chinese version of the Structured Clinical Interview for DSM-IV (SCID-IV) (Phillips & Liu, 2011) was administered by a trained psychiatrist, a detailed history of medication use was obtained, and the inclusion and exclusion criteria for the study were assessed. Enrollment criteria for UT included: 1) meets DSM-IV diagnostic

criteria for schizophrenia, and 2) completely untreated (had never received any antipsychotic treatment) or virtually untreated (less than 15 days of lifetime antipsychotic treatment, no antipsychotic treatment in the last 5 years, and no history of hospitalization in a psychiatric ward). Criteria for TC included: 1) meets DSM-IV diagnostic criteria for schizophrenia, 2) received antipsychotic medications regularly (i.e., for more days than not) for at least 30 days at some point during the first two years after the initial onset of psychotic symptoms, 3) received antipsychotic medications regularly over the 90 days before assessment, and 4) had similar duration of illness and demographic factors (i.e., gender, age, years of schooling, urban v. rural residence, ethnicity) to UT. Criteria for HC included: 1) never had psychotic symptoms, 2) never received antipsychotic treatment, 3) not receiving treatment for any mental disorder at the time of assessment, and 4) had similar demographic characteristics to UT.

Exclusion criteria for all three groups included: 1) less than 18 years of age; 2) history of intellectual disability starting in childhood, epilepsy, severe brain injury, or other organic mental disorder; 3) current substance use disorder; 4) currently suicidal or violent; and 5) (for UT and TC) severe psychotic symptoms that precluded completion of the cognitive assessment (i.e., score of '7' on the Positive and Negative Symptom Scale [PANSS; Kay, Fiszbein, & Opler, 1987] items assessing delusions, hallucinations, or conceptual disorganization).

4.2.2 Assessments

4.2.2.1 Demographic information

Demographic variables collected included age, years of schooling, gender, ethnicity (i.e., Han [the majority ethnic group in China] v. any minority ethnic group), and location of residence (urban v. rural).

4.2.2.2 Diagnosis and duration of illness

Participants were diagnosed by trained psychiatrists using the research version of SCID-IV. The DSM-IV diagnoses were converted to DSM-5 diagnoses by listening to the audiotapes of the diagnostic interviews; all individuals meeting DSM-IV criteria for schizophrenia also met DSM-5 criteria. Duration of illness was the time interval between the first onset of hallucinations, delusions, or disorganized thinking and the interview.

4.2.2.3 Social cognition

The 70-item RMET-CV-R used to assess social cognition employed the same black-and-white stimulus pictures (15 cm * 6cm) as the 70-item RMET-CV, 36 of which were identical to those used in the original English version of RMET. Pictures are presented on a computer screen, and participants first select one of four adjectives (shown with the picture) that best describes the emotion depicted and then judge the gender of the depicted person. There is no time limit for each selection. When participants find it hard to decide, they are asked to make their best guess. Before starting the formal test, there is a training session using two pictures of Asian eyes.

In this study 59.4% of the sample (n=384) were over 60 years of age or had less than 5 years of schooling; that is, they did not meet the inclusion criteria used in China to develop norms for the MCCB (limited to individuals <60 years of age who had $\geq$5 years of full-time schooling [Shi et al., 2015]). Moreover, 88% of respondents had never used a computer. To conduct the test with these respondents, it was necessary to make several changes in the administration of the test (without changing the presented items or optional descriptors provided in the original versions of the test). We trained respondents to use the computer mouse, provided more extensive training about the task's requirements, and had interviewers read test items to illiterate or semi-literate respondents. Details of these changes are provided

in the Supplementary Materials.

4.2.2.4 Other instruments

In addition to the RMET-CV-R, the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975) was administered to all respondents, and PANSS was administered to respondents with schizophrenia.

4.2.3 Outcome measures

4.2.3.1 RMET-CV-R completion status

Given the possibility that some older or uneducated respondents could not successfully complete RMET-CV-R, we classified the completion status of each test administration as 'successfully completed' or 'incomplete' (subclassified as incomplete due to 'refusal to do the test', 'inability to complete the test because of difficulty understanding the provided emotional descriptors', or 'inability to understand what is expected in the test'). Operational criteria for these classifications are provided in the Supplementary Materials.

4.2.3.2 RMET-CV-R test scores

The RMET-CV-R test score recorded for all successfully completed tests is the percentage of correct judgments about the emotion exhibited in the presented slides (i.e., number of correct responses divided by 70).

4.2.4 Modelling strategy

- **Characteristics of participants**

All comparisons between UT, TC, and HC groups were conducted on the full sample and separately on four subsamples. To facilitate comparison of results

between respondents who could complete the computer-based test on their own with those who had the test read by the interviewer (a new method of administering the test), we stratified the sample into ‘self-completed’ and ‘interviewer-completed’ subsamples. To facilitate comparison of results between respondents typically included in previous studies with RMET with those of older, less educated respondents (typically excluded from such studies), we used the inclusion criteria for the Chinese MCCB normalization sample for RMET (Shi et al., 2015) to stratify the sample into an ‘educated-younger’ subsample (of schooling). The distribution of respondents between the three groups in the full sample and the 2 subsamples are shown in Figure 4.1.

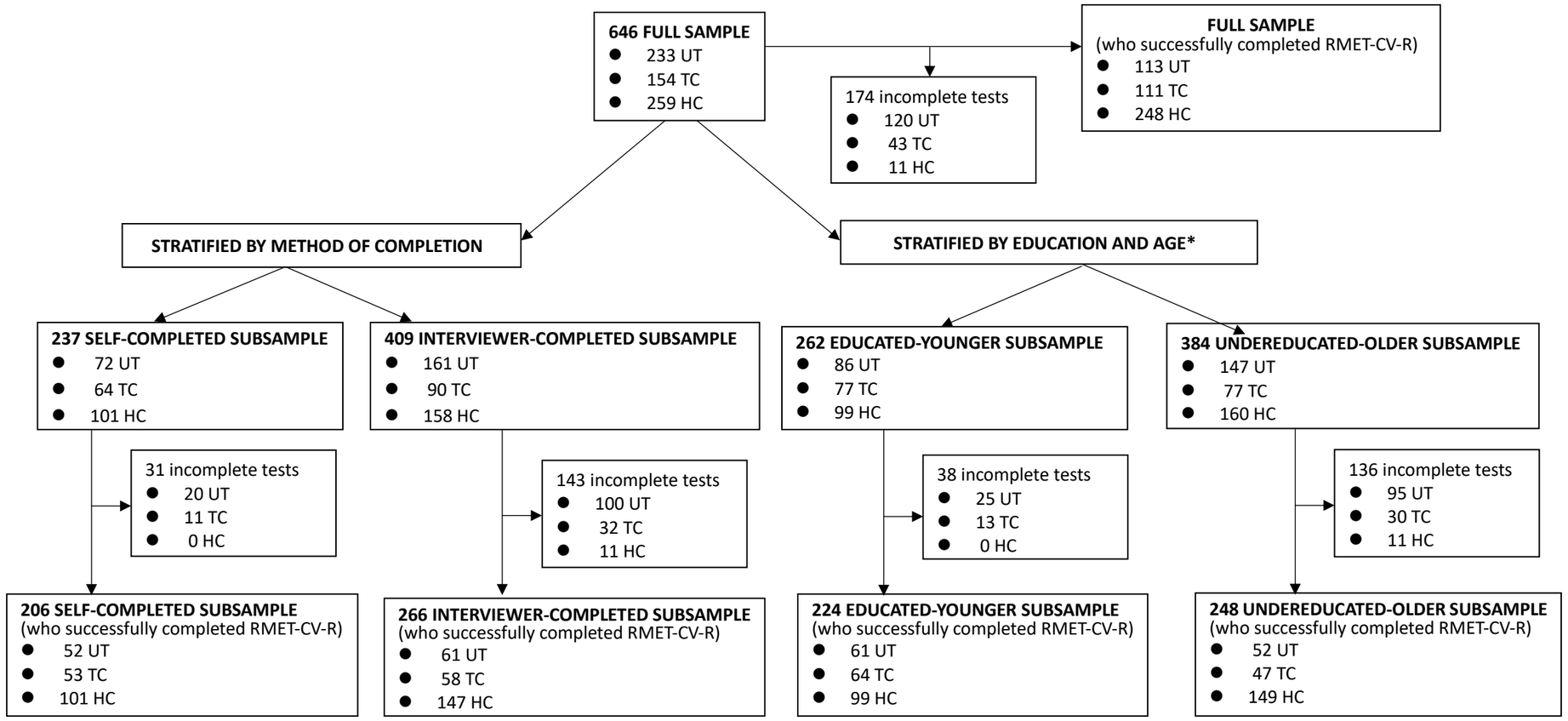

Figure 4.1 Enrollment, stratification and completion status of participants

UT, Untreated case; TC, Treated Control; HC, Healthy Control

*respondents with >5 years of schooling and <60 are classified as 'educated-younger' and those with <5 years of schooling or ≥60 as 'undereducated-older'

- **Completion status of RMET-CV-R**

Baseline covariates between the study groups (shown in Table 4.1) were compared using t-tests (2 groups) or ANOVA F-tests (3 groups) for continuous variables, Mann-Whitney tests (2 groups) or Kruskal-Wallis tests (3 groups) for ordinal variables, and Chi-square tests for categorical variables. The Dunn Test (Dunn, O.J., 1961; Dinno, A., 2015) is used for pairwise multiple comparisons of ranked results and a Tukey-type multiple comparison test was used for pairwise comparisons of dichotomous results (Zar, 1999).

Comparison of the successful completion rates between groups are assessed using Chi square (Table 4.2). Multiple logistic binary regression was firstly used to compare rates of successful completion of RMET-CV-R among groups while adjusting for the linear terms of demographic covariates (i.e., age, gender, years of schooling, residence, and ethnicity) using equation one (Table 4.3).

$$logit[Pr(y_i = 1|X_i)] = X_i'\beta + \varepsilon_i \quad (1)$$

X_i included seven variables, two dummy variables of group assignment, age in years, gender (1=female, 2=male), years of schooling, residence (1=urban, 2=rural), and ethnicity (1=Han, 2=ethnic minority group).

In the analyses of the educated-younger subsample, since 100% of respondents in the HC group successfully completed RMET-CV-R, Firth's penalized likelihood method (Firth, 1993) was employed to estimate the parameters to reduce potential bias caused by data separation.

To capture the potential probability that years of schooling might have different effects on RMET scores in three groups and the relationship between years of schooling and RMET score might not be linear, two interaction terms were added to the logistic regression in the whole sample, group dummy * years of schooling and the quadratic term of years of schooling (shown in equation two, where D_i denoted the group dummy variable and X_i denoted same as it in equation one).

$$logit[Pr(y_i = 1|X_i)] = X_i'\beta + \gamma D_i * edu_i + \delta edu_i^2 + \varepsilon_i \quad (2)$$

Logistic regression was also conducted only in two groups of patients and included duration of illness (in years) with/without PANSS score.

- **RMET-CV-R scores using linear regression**

Crude RMET-CV-R scores between groups are compared using ANOVA F-tests (Table 4.5).

Two linear-regression based methods were used to compare RMET-CV-R scores between groups while adjusting for demographic covariates and controlling heteroscedasticity. Confidence intervals of coefficients were calculated using the robust standard error (Table 4.6).

The first analysis, restricted to respondents who successfully completed the RMET-CV-R, used generalized least square (GLS) linear regression to estimate the regression parameters (equation 3) separately in the whole sample and sub-samples, where X_i denoted same as it in equation one.

$$RMET_i = X_i'\theta + \varepsilon_i \quad (3)$$

Also, to capture the potential probability that years of schooling might have variant effects at different level of educational attainment (e.g., the increase of educational attainment might have a more significant effect in the lower-educated group) or at different age groups, a quadratic term of years of schooling and an interaction term between age and years of schooling were added into equation three (shown in equation four).

$$RMET_i = X_i'\theta + \zeta edu_i^2 + \lambda edu_i * age_i + \varepsilon_i \quad (4)$$

- **RMET-CV-R scores using Heckman estimation**

The second analysis included all subjects in the sample and used Heckman's method (Heckman, 1976) to adjust for potential selection bias. The rate of successful completion was 48.5% in UT, 70.1% in TC, and 95.6% in HC, respectively. The failure to successfully complete RMET-CV-R is presumably related to different levels of

overall cognitive functioning, so it was important to make a further adjustment of the RMET-CV-R scores to correct for the non-random selection bias introduced by differential rates of non-completion across the three study groups. The Heckman approach adjusts estimates of the factors associated with the RMET-CV-R score based on the characteristics of individuals who did and did not successfully complete RMET-CV-R.

Heckman estimation involved two equations: a probit model for the probability of successfully completing RMET-CV-R and a linear model for the RMET-CV-R score conditional on those who successfully complete RMET-CV-R.

Let equation that determined the sample selection (i.e., whether or complete the test) be

$$z_i^* = W'\xi + u_i \qquad (5)$$

and the equation of primary interest, i.e., RMET score, still be

$$RMET_i = X_i'\theta + \varepsilon_i. \qquad (6)$$

The sampling rule was that RMET score was observed only when the latent variable z_i^* was greater than zero. The conditional mean that applied to the observed samples was

$$E[RMET|X, z_i^* > 0] = X_i'\theta + \mathrm{E}[\varepsilon_i|u_i > -W'\xi]. \qquad (7)$$

When the error terms of u_i and ε_i had a bivariate normal distribution with zero means and correlation ρ,

$$E[RMET|X, z_i^* > 0] = X_i'\theta + \rho\sigma_\varepsilon\lambda(\alpha_u), \qquad (8)$$

where $\lambda(\alpha_u) = \varphi(W'\xi/\sigma_u)/\Phi(W'\xi/\sigma_u)$ and a positive ρ in this study indicated that participants who successfully complete the RMET-CV-R test had better social cognition than those who could not successfully complete it.

So,

$$RMET_i|z_i^* > 0 = X_i'\theta + \beta_\lambda\lambda(W'\xi/\sigma_u) + v_i. \qquad (9)$$

Therefore, when the $\beta_\lambda \lambda(W'\xi/\sigma_u)$ was omitted in the GLS regression and only X_i' was included, the equation produced inconsistent estimation of θ.

The Heckman model can be estimated either by a two-step estimation procedure or a single-step full information maximum likelihood (FIML) model (Greene WH, 2018). The FIML model is used in this study because it is more efficient (Nawata, 1994; Greene WH, 2018). The full log-likelihood function was shown below (equation 10):

$$lnL = \sum_{z=1} ln\left[\frac{\exp\left(-\left(\frac{1}{2}\right)\varepsilon_i^2/\sigma_\varepsilon^2\right)}{\sigma_\varepsilon\sqrt{2\pi}}\Phi\left(\frac{\rho\varepsilon_i/\sigma_\varepsilon + W_i'\xi}{\sqrt{1-\rho^2}}\right)\right] + \sum_{z=0}\ln\left[1-\Phi(W_i'\xi)\right]. \quad (10)$$

Group assignment (untreated patients, age, gender, years of schooling, residence, ethnicity, the quadric term of years of schooling, an interaction term between years of schooling, and an interaction term between group assignment and years of schooling were in included in the vector W_i' and assumed to be related to the RMET completion status, and group assignment (untreated patients, age, gender, years of schooling, residence, ethnicity, the quadric term of years of schooling, and an interaction term between years of schooling were included in the vector X_i'.

All tests were two-tailed and the level of significance was set at 0.05. All analysis was conducted using STATA (version 17 SE).

4.3 Results

4.3.1 Characteristics of participants

Participants had a mean (sd) age of 51.1 (11.7) years, a median (IQR) duration of formal schooling of 4 (0-7) years (195 [30.2%] had never attended school), 55.7% were female, 90.9% were living in rural communities, 40.7% were members of a minority ethnic group, and their mean MMSE total score was 22.0 (10.7). The full sample was stratified into self-completed (n=237, 36.7%) and interviewer-completed (n= 409, 63.3%) subsamples, and into educated-younger (n=262, 40.6%) and undereducated-older (n=384, 59.4%) subsamples.

TC were younger than UT and HC in the full sample and all four subsamples (Table 4.1), though the differences were not statistically significant in all cases. With the single exception of a higher proportion of female TC in the undereducated-older subsample, gender, years of schooling, residence, and ethnicity were not significantly different between UT, TC, and HC in the full sample or any of the subsamples. In the full sample and all subsamples, the MMSE total score was significantly greater in HC than in TC and significantly greater in TC than in UT, even after adjusting for demographic factors.

Among the 387 participants with schizophrenia, mean duration of illness was 22.5 (11.6) years and mean PANSS total score was 71.7 (22.1). Duration of illness was significantly greater in UT than in TC in the full sample and in the interviewer-completed and undereducated-older subsamples. In the full sample and all four subsamples, PANSS total score was significantly greater in UT than in TC (Table 4.1), even after adjusting for demographic factors. In the TC group, patients used antipsychotic medications for a mean of 77.2% (28.6%) of the total duration of illness.

Table 4.1 Demographic characteristics of the full sample and different subsamples

Characteristics	Full sample			Respondents assigned to self-complete RMET-CV-R v. Respondents assigned to have RMET-CV-R administered by interviewer						Respondents with ≥ 5 years of schooling and <60 ('educated-younger') v. respondents with <5 years of schooling or ≥ 60 ('undereducated-older')					
				self-completed subsample			interviewer-completed subsample			educated-younger subsample			undereducated-older subsample		
	UT n=233	TC n=154	HC n=259	UT n=72	TC n=64	HC n=101	UT n=161	TC n=90	HC n=158	UT n=86	TC n=77	HC n=99	UT n=147	TC n=77	HC n=160
Age in years, mean (sd)[a]	52.6 (12.4)	48.1 (9.6)	51.7 (11.8)	44.8 (10.3)	42.5 (7.0)	47.2 (11.2)	56.1 (11.8)	52.0 (9.3)	54.5 (11.2)	45.0 (9.0)	43.0 (6.4)	45.3 (9.0)	57.0 (12.1)	53.1 (9.7)	55.6 (11.6)
	F=7.62, p<0.001 [UT,HC>TC]			**F=4.44, p=0.013 [HC>TC]**			**F=3.94, p=0.020 [UT>TC]**			F=1.81, p=0.165 [ns]			**F=3.03, p=0.049 [UT>TC]**		
Years of schooling, median(IQR)[b]	3 (0-6)	5 (0-8)	4 (0-7)	7 (5-9)	8 (6-9)	8 (5-8)	1 (0-3)	1.5 (0-4)	0 (0-3)	7 (5-8)	8 (5-9)	8 (5-9)	0 (0-3)	0 (0-3)	0.5 (0-3)
	Chi^2=5.64, p=0.060 [ns]			Chi^2=2.15, p=0.327 [ns]			Chi^2=2.04, p=0.361 [ns]			Chi^2=1.10, p=0.554 [ns]			Chi^2=0.50, p=0.779 [ns]		
Gender: female (%)[c]	127 (54.5%)	97 (63.0%)	136 (52.5%)	32 (44.4%)	31 (48.4%)	41 (40.6%)	95 (59.0%)	66 (73.3%)	95 (60.1%)	37 (43.0%)	39 (50.7%)	43 (43.4%)	90 (61.2%)	58 (75.3%)	93 (58.1%)
	Chi^2=4.52, p=0.104 [ns]			Chi^2=0.99, p=0.609 [ns]			Chi^2=5.73, p=0.057 [ns]			Chi^2=1.21, p=0.547 [ns]			**Chi^2=6.82, p=0.033, [TC>HC]**		
Residence: rural (%)[c]	213 (91.4%)	138 (89.6%)	236 (91.1%)	64 (88.9%)	53 (82.8%)	89 (88.1%)	149 (92.6%)	85 (94.4%)	147 (93.0%)	78 (90.7%)	65 (84.4%)	87 (87.9%)	135 (91.8%)	73 (94.8%)	149 (93.1%)
	Chi^2=0.40, p=0.820 [ns]			Chi^2=1.32, p=0.516 [ns]			Chi^2=0.33, p=0.847 [ns]			Chi^2=1.50, p=0.473 [ns]			Chi^2=0.69, p=0.708 [ns]		
Ethnicity: Han (%)[c]	133 (57.1%)	100 (64.9%)	150 (57.9%)	50 (69.4%)	49 (76.6%)	72 (71.3%)	83 (51.6%)	51 (56.7%)	78 (49.4%)	60 (69.8%)	59 (76.6%)	70 (70.7%)	73 (49.7%)	41 (53.2%)	80 (50.0%)
	Chi^2=2.71, p=0.258 [ns]			Chi^2=0.92, p=0.631 [ns]			Chi^2=1.23, p=0.540 [ns]			Chi^2=1.11, p=0.573 [ns]			Chi^2=0.29, p=0.865 [ns]		
Mini Mental Status Examination (MMSE): mean total score (sd)[a,d]	16.9 (10.0)	22.9 (9.5)	26.6 (7.6)	24.6 (9.0)	29.5 (7.0)	32.3 (4.2)	13.4 (8.4)	18.1 (8.2)	22.9 (7.0)	23.5 (9.2)	29.0 (6.7)	32.1 (4.8)	13.0 (8.3)	16.6 (7.8)	23.2 (7.1)
	F=68.99, p<0.001 [HC>TC>UT]			**F=27.38, p<0.001 [HC>TC>UT]**			**F=56.47, p<0.001 [HC>TC>UT]**			**F=34.59, p<0.001 [HC>TC>UT]**			**F=66.83, p<0.001 [HC>TC>UT]**		
Mean (sd) duration of illness in months[b]	286 (153)	245 (109)	----	232 (139)	214 (88)	----	311 (153)	267 (117)	----	221 (121)	223 (82)	----	324 (157)	266 (127)	----
	Z=2.43, p=0.016 [UT>TC]			Z=0.30, p=0.765 [ns]			**Z=2.05, p=0.041 [UT>TC]**			Z=-0.37, p=0.712 [ns]			**Z=2.44, p=0.015 [UT>TC]**		
Mean (sd) PANSS total score[e]	79.0 (20.3)	60.7 (20.3)	----	73.4 (16.8)	55.8 (18.6)	----	81.4 (21.2)	64.3 (20.7)	----	74.7 (17.6)	56.8 (17.9)	----	81.4 (21.3)	64.7 (21.8)	----
	t=8.65, p<0.001 [UT>TC]			**t=5.79, p<0.001 [UT>TC]**			**t=6.20, p<0.001 [UT>TC]**			**t=6.44, p<0.001 [UT>TC]**			**t=5.52, p<0.001 [UT>TC]**		

RMET-CV-R Revised Chinese Version of the Reading the Mind in the Eyes; UT, untreated group; TC, treated controls; HC, healthy controls; ns, not significant
statistically significant results are **bolded**

[a] ANOVA is used to compare means across groups and Tukey test is use for pairwise comparisons

[b] Kruskal-Wallis test is used to compare non-normal continuous variables in three groups and Mann-Whitney test is used to compare non-normal continuous variables in two groups.

[c] If the overall Chi-square value is statistically significant, pairwise comparisons between groups are made using a Tukey-type test for proportions (Zar, 1999)

[d] 13 UT (3 in the educated-younger subsample and 10 in the undereducated-older subsamples) and 1 TC in the undereducated-older subsample did not complete MMSE.

[e] Student's t test is used to compare the mean of continuous variables between untreated patients and treated controls

4.3.2 Completion status of RMET-CV-R

Among the 646 participants, 472 (73.1%) successfully completed RMET-CV-R. The 174 (26.9%) who did not successfully complete the test included 37 (5.7%) who refused, 78 (12.1%) who understood what was required but were unable to complete the test (primarily due to limited vocabulary), and 59 (9.1%) who were unable to understand test instructions despite several attempts to explain the requirements.

In the full sample and all subsamples, HC had the highest successful completion rates, TC had intermediate successful completion rates, and UT had the lowest successful completion rates (Table 4.2). Differences between HC and the two patient groups were statistically significant in the full sample and all subsamples, but the difference between TC and UT was only statistically significant in the full sample and in the interviewer-completed and undereducated-older subsamples.

Results of multivariate binary logistic regression analysis (Table 4.3) of the full sample show that group status (HC>TC>UT), younger age, male gender, and higher levels of education are all significantly associated with successful completion of RMET-CV-R, but residence and ethnicity are not. The same pattern is evident in the interviewer-completed and undereducated-older subsamples. However, in the self-completed and educated-younger subsamples, while successful completion rates remain significantly higher in HC than in UT and TC, the higher successful completion rates in TC than in UT are not statistically significant, and successful completion is not significantly different by age, gender, or years of schooling.

As shown in Panela A and B of Figure 4.2, in the full sample and all subsamples after adjusting for inter-group differences in age, gender, education, ethnicity, and residence, the percent successful completion is highest in HC, intermediate in TC, and least in UT. However, the magnitude of the difference in successful completion rates between the self-completed versus interviewer-completed subsamples and between educated-younger versus undereducated-older subsamples is smallest in HC, intermediate in TC, and largest in UT.

Table 4.2 RMET-CV-R completion status in the full sample and different subsamples

	Full sample			**Respondents assigned to self-complete RMET-CV-R (self-completed subsample)**			**Respondents assigned to have RMET-CV-R administered by interviewer (interviewer-completed subsample)**			**Respondents with ≥ 5 years of schooling and <60 (educated-younger subsample)**			**Respondents with <5 years of schooling or≥60 (undereducated-older subsample)**		
	UT	TC	HC	UT	TC	HC	UT	TC	HC	UT	TC	HC	UT	TC	HC
COMPLETION STATUS[a]	n=233	n=154	n=259	n=72	n=64	n=101	n=161	n=90	n=158	n=86	n=77	n=99	n=147	n=77	n=160
INCOMPLETE TEST	120 (51.5%)	43 (27.9%)	11 (4.2%)	20 (27.8%)	11 (17.2%)	0 (0.0%)	100 (62.1%)	32 (35.6%)	11 (7.0%)	25 (29.1%)	13 (16.9%)	0 (0.0%)	95 (64.6%)	30 (39.0%)	11 (6.9%)
Refused to complete test	*25 (10.7%)*	*9 (5.8%)*	*3 (1.2%)*	*7 (9.7%)*	*5 (7.8%)*	*0 (0.0%)*	*18 (11.2%)*	*4 (4.4%)*	*3 (1.9%)*	*8 (9.3%)*	*5 (6.5%)*	*0 (0.0%)*	*17 (11.6%)*	*4 (5.2%)*	*3 (1.9%)*
Unable to complete test	*47 (20.2%)*	*24 (15.6%)*	*7 (2.7%)*	*2 (2.8%)*	*2 (3.1%)*	*0 (0.0%)*	*45 (28.0%)*	*22 (24.4%)*	*7 (4.4%)*	*6 (7.0%)*	*4 (5.2%)*	*0 (0.0%)*	*41 (27.9%)*	*20 (26.0%)*	*7 (4.4%)*
Did not understand test	*48 (20.6%)*	*10 (6.5%)*	*1 (0.4%)*	*11 (15.3%)*	*4 (6.3%)*	*0 (0.0%)*	*37 (23.0%)*	*6 (6.7%)*	*1 (0.6%)*	*11 (12.8%)*	*4 (5.2%)*	*0 (0.0%)*	*37 (25.2%)*	*6 (7.8%)*	*1 (0.6%)*
SUCCESSFULLY COMPLETED TEST	113 (48.5%)	111 (72.0%)	248 (95.8%)	52 (72.2%)	53 (82.8%)	101 (100.0%)	61 (37.9%)	58 (64.4%)	147 (93.0%)	61 (70.9%)	64 (83.1%)	99 (100.0%)	52 (35.4%)	47 (61.0%)	149 (93.1%)
	Chi^2=139.27, p<0.001 [HC>TC>UT]			**Chi^2=29.83, p<0.001 [HC>UT, TC]**			**Chi^2=106.68, p<0.001 [HC>TC>UT]**			**Chi^2=31.86, p<0.001 [HC>TC,UT]**			**Chi^2=112.4, p<0.001 [HC>TC>UT]**		
METHOD OF COMPLETING TEST IN PERSONS WHO SUCCESSFULLY COMPLETE TEST	n=113	n=111	n=248	n=52	n=53	n=101	n=61	n=58	n=147	n=61	n=64	n=99	n=52	n=47	n=149
Self-completed	52 (46.0%)	53 (47.8%)	101 (40.7%)	52 (100.0%)	53 (100.0%)	101 (100.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	51 (83.6%)	52 (81.3%)	90 (90.9%)	1 (1.9%)	1 (2.1%)	11 (7.4%)
Interviewer-completed	61 (54.0%)	58 (52.2%)	147 (59.3%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	61 (100.0%)	58 (100.0%)	147 (100.0%)	10 (16.4%)	12 (18.8%)	9 (9.1%)	51 (98.1%)	46 (97.9%)	138 (92.6%)
	Chi^2=1.88, p=0.391 [ns]			-----			----			Chi^2=3.50, p=0.174 [ns]			Chi^2=3.45, p=0.179 [ns]		

RMET-CV-R Revised Chinese Version of the Reading the Mind in the Eyes; UT, untreated group; TC, treated controls; HC, healthy controls; ns, not significant

statistically significant results are **bolded**

[a] Chi square compares proportion of completed tests across three groups (df=1), if statistically significant, pairwise comparisons between groups are made using a Tukey-type test for proportions (Zar, 1999)

Table 4.3 Multivariate logistic regression comparing rates of successful completion of RMET-CV-R between UT, TC, and HC in the full sample and two subsamples

		Full sample n=646		Respondents assigned to self-complete RMET-CV-R (self-completed subsample) n=236[a]		Respondents assigned to have RMET-CV-R administered by interviewer (interviewer-completed subsample) n=407[b]		Respondents with ≥ 5 years of schooling and <60 (educated-younger subsample) n=262		Respondents with <5 years of schooling or ≥ 60 (undereducated-older subsample) n=384	
		OR	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**
Group	TC v. UT*	**2.73**	**1.64-4.56**	1.60	0.68-3.77	3.41	1.85-6.30	1.84	0.85-3.97	**3.49**	**1.81-6.74**
	HC v. UT*	**34.40**	**17.79-66.50**	**85.54**	**5.03-1455.53**	**32.07**	**15.71-65.46**	**84.93**	**5.06-1426.65**	**33.13**	**16.01-68.58**
	HC v. TC*	**12.59**	**6.07-26.11**	**53.40**	**3.01-946.95**	**9.40**	**4.24-20.82**	**46.26**	**2.68-798.87**	**9.50**	**4.20-21.49**
Age in years		**0.97**	**0.95-0.99**	0.97	0.93-1.02	0.97	0.95-0.99	0.98	0.93-1.03	**0.97**	**0.94-0.99**
Gender (1=female*, 2=male)		**1.83**	**1.16-2.88**	2.06	0.89-4.78	1.71	0.99-2.96	1.56	0.74-3.29	**1.90**	**1.04-3.48**
Years of schooling		**1.25**	**1.15-1.35**	**1.27**	**1.00-1.61**	1.28	1.13-1.45	1.22	0.99-1.49	**1.36**	**1.16-1.59**
Residence (1=urban*, 2=rural)		0.44	0.16-1.23	0.94	0.23-3.91	0.30	0.09-1.02	0.64	0.15-2.80	0.37	0.10-1.35
Ethnicity (1=Han*, 2=minority group)		1.38	0.87-2.18	1.17	0.46-2.99	1.41	0.83-2.38	1.41	0.60-3.32	1.32	0.76-2.29
Adjusted percent (95% CI) of successful completion of RMET-CV-R computed from regression											
UT		50.4% (44.7%-56.0%)		73.9% (62.2%-83.1%)		37.3% (30.4%-44.3%)		72.1% (61.5%-80.7%)		35.8% (28.6%-43.1%)	
TC		69.5% (62.2%-76.8%)		81.9% (69.6%-89.9%)		63.2% (53.2%-73.1%)		82.6% (72.1%-89.8%)		62.0% (51.3%-72.7%)	
HC		95.7% (93.3%-98.1%)		99.6% (93.6%-100.0%)		93.1% (89.3%-96.9%)		99.5% (93.1%-100.0%)		92.7% (88.7%-96.7%)	

RMET-CV-R Revised Chinese Version of the Reading the Mind in the Eyes; TC, Treated Control; UT, Untreated case; HC, Healthy Control
statistically significant results are **bolded**
* denotes the reference group
[a] using firth logit regression because 100% of the participants in HC group completed the test

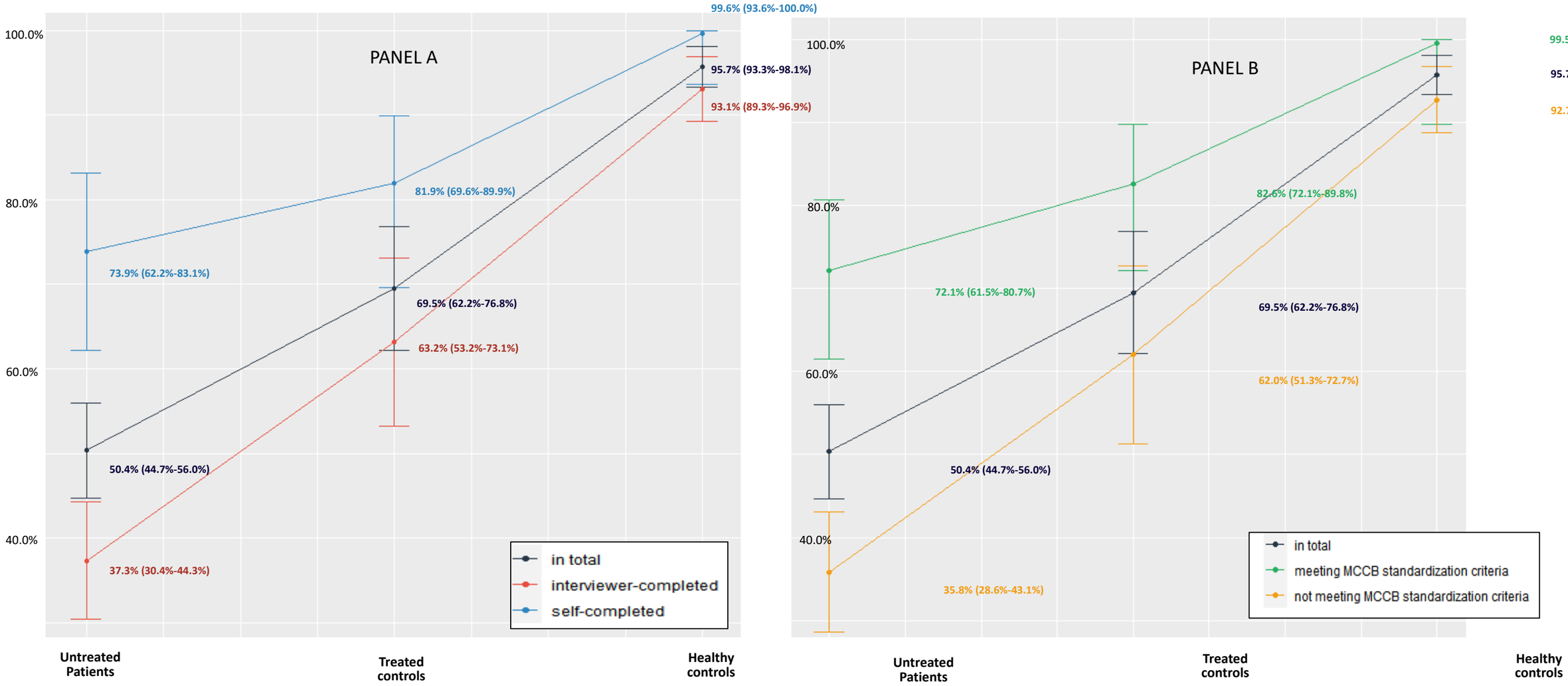

Figure 4.2 Adjusted percent (and 95% CI) of untreated cases, treated controls and healthy controls who successfully complete RMET-CV-R in the self-completed and interview-completed subgroups (Panel A) and in respondents who do or do not meet MCCB standardization criteria (Panel B);

As shown in Supplement Table 4.1A and 4.1B, when limiting analyses to the two patient groups and controlling for duration of illness, adjusted percent successful completion in TC is significantly higher than that in UT (69.0% v. 50.8%, OR=2.61, 95% CI=1.58-4.29); however, when controlling for both duration of illness and PANSS total score, this difference is no longer statistically significant (62.1% v. 55.4%, OR=1.47, 95% CI=0.85-2.56).

By including the quadratic term of years of schooling and an interaction term between group assignments, equation two captured significant none linear relationship between years of schooling and RMET score, especially in HC group. As shown in Figure 3, the marginal effect of years of schooling gradually decreased with the increase of years of schooling in UT and TC group, and it declined dramatically in HC group when the years of schooling went up from zero to three but were almost zero when years of schooling exceeded three years. In the none-linear regression, group assignment, age and gender still remained significant.

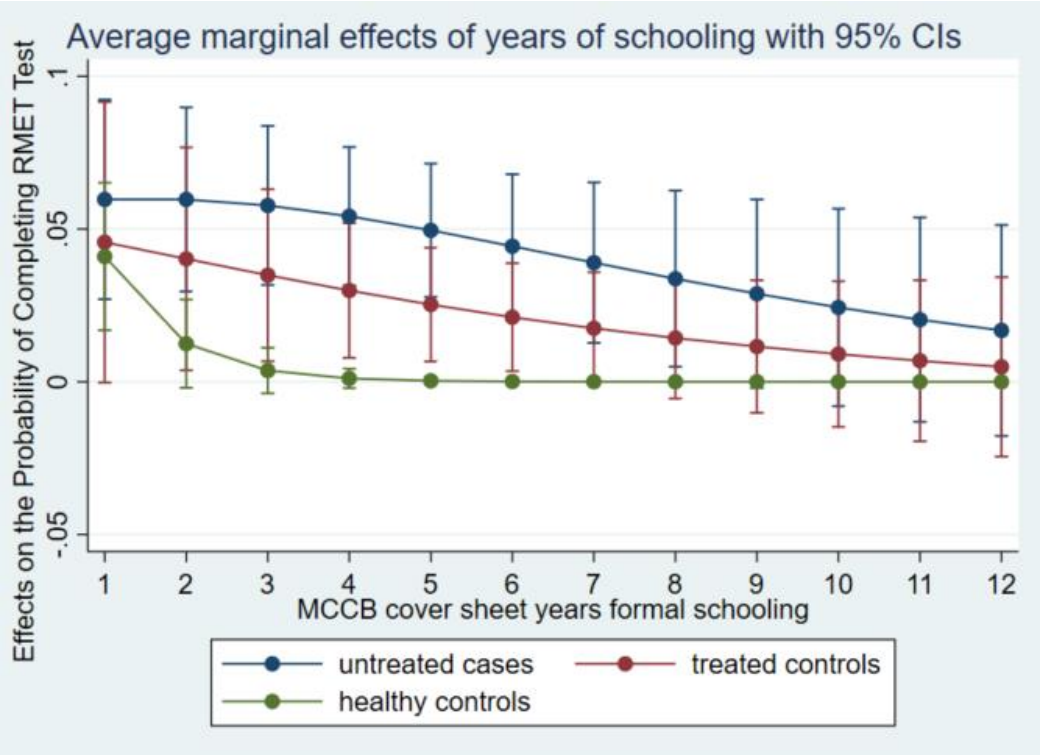

Figure 4.3 Marginal effects (and 95% CI) of years of schooling on the probability of completing RMET-CV-R test in three groups of participants

Table 4.4 Multivariate logistic regression comparing rates of successful completion of RMET-CV-R between UT, TC, and HC in the full sample including quadratic term of years of schooling and the interaction term of group and years of schooling

		Full sample n=646			
		Coefficient	**P value**	**OR**	**95% CI**
Group (untreated group as reference group)	TC v. UT*	**1.23**	**0.001**	**3.44**	**1.69-7.00**
	HC v. UT*	**3.12**	**<0.001**	**22.66**	**10.28-49.95**
Age in years		**-0.03**	**0.005**	**0.97**	**0.95-0.99**
Gender (1=female*, 2=male)		**0.58**	**0.015**	**1.78**	**1.12-2.88**
Years of schooling		**0.29**	**0.003**	**1.25**	**1.15-1.35**
Quadratic term of years of schooling		-0.01	0.428	0.99	0.97-1.01
Residence (1=urban*, 2=rural)		-0.90	0.078	0.41	0.15-1.11
Ethnicity (1=Han*, 2=minority group)		0.30	0.201	1.35	0.85-2.14
Interaction: group dummy*years of schooling					
TC		-0.07	0.332	0.93	0.81-1.07
HC		**0.97**	**0.043**	**2.64**	**1.03-6.78**
Adjusted percent (95% CI) of successful completion of RMET-CV-R computed from regression					
UT		50.4% (44.8%-56.0%)			
TC		69.5% (62.2%-76.8%)			
HC		95.7% (93.3%-98.1%)			

RMET-CV-R Revised Chinese Version of the Reading the Mind in the Eyes; TC, Treated Control; UT, Untreated case; HC, Healthy Control
statistically significant results are **bolded**
* denotes the reference group

4.3.3 RMET-CV-R test scores

Among the 472 successfully completed tests, the mean (sd) completion time was 14.4 (6.1) minutes, and the mean RMET-CV-R score was 42.1% (16.1%) for the recognition of the emotion expressed in the 70 presented pictures. Respondents correctly identified gender in 81.3% (13.1%) of the pictures, which indicates that they were attending to the task.

In the full sample and all subsamples, the recognition of emotion was greatest in

HC, intermediate in TC, and lowest in UT (Table 4.5); however, these differences were not statistically significant in all cases. For all three groups (UT, TC, and HC), emotion recognition was significantly greater in self-completed versus interviewer-completed subsamples and significantly greater in educated-younger versus undereducated-older subsamples. Moreover, the magnitude of the difference in emotion recognition scores between HC and UT was greater in the self-completed subsample (13.6%) than in the interviewer-completed subsample (7.2%) and greater in the educated-younger subsample (13.8%) than in the undereducated-older subsample (7.9%).

Results of the multivariate linear regression (Table 4.6) show that after adjusting for demographic covariates, HC had significantly higher RMET-CV-R scores than TC and UT both in the full sample and in all four subsamples. TC also had higher scores than UT in the full sample and in all four subsamples, but the difference was not statistically significant in the interviewer-completed and undereducated-older subsamples.

As shown in panels A and B of Figure 4.4, the adjusted mean RMET-CV-R scores in the full sample and all four subsamples are highest in HC, intermediate in TC, and lowest in UT. However, differences in the adjusted mean scores between UT and TC are greater in self-completed (7.3%) versus interviewer-completed (1.3%) subsamples and greater in educated-younger (5.7%) versus undereducated-older (2.1%) subsamples.

As shown in Supplement Table 4.2A and 4.2B, when limiting analyses to the two patient groups, after adjusting for demographic variables, except for the significant negative relationship between duration of illness and RMET score in the undereducated-older subsample, duration of illness and PANSS total score were *not* significantly associated with patients' RMET-CV-R scores.

The non-linear model including the quadratic term of years of schooling found that the marginal effect of years of schooling decreased with the increase of years of schooling (Figure 4.5 Panel A), i.e., years of schooling had a more significant effect on the lower educated participants. The interaction term between age and years of

schooling was also significant, and years of schooling had a stronger effect when age increased. Group differences remained significant in the non-linear model, and younger age, male, and urban residence were significantly related to higher RMET scores in the full sample.

Table 4.5 Unadjusted RMET-CV-R mean (sd) scores (percent correct responses in 70 pictures) for UT, TC and HC in all respondents who successfully completed RMET-CV-R and in different subsamples of respondents who successfully completed RMET-CV-R [a]

	All respondents who successfully completed RMET-CV-R			Respondents who successfully self-completed RMET-CV-R (self-completed subsample)			Respondents who successfully completed RMET-CV-R administered RMET-CV-R by interviewers (interviewer-completed subsample)			Respondents who successfully completed RMET-CV-R with ≥5 years of schooling and <60 (educated-younger subsample)			Respondents who successfully completed RMET-CV-R with <5 years of schooling or ≥ 60 (undereducated-older subsample)		
	UT n=113	TC n=111	HC n=248	UT n=52	TC n=53	HC n=101	UT n=61	TC n=58	HC n=147	UT n=61	TC n=64	HC n=99	UT n=52	TC n=47	HC n=149
RMET-CV-R emotion score	36.4% (14.3%)	40.5% (14.9%)	43.8% (17.0%)	42.2% (15.1%)	49.9% (12.2%)	54.9% (12.4%)	31.4% (11.5%)	31.9% (11.8%)	36.2% (15.4%)	41.1% (14.6%)	47.3% (13.1%)	54.5% (13.2%)	30.8% (11.8%)	31.2% (12.1%)	36.7% (15.3%)
	F=8.66, p<0.001 [HC>UT]			**F=16.29, p<0.001 [HC,TC>UT]**			F=3.54, p=0.030 [ns]			**F=19.00, p<0.001 [HC>TC>UT]**			**F=4.85, p=0.009 [HC>UT]**		
				COMPARISION OF RESULTS FOR EACH GROUP BETWEEN SELF-COMPLETION AND INTERVIEWER-COMPLETION RESPONDENTS **UT: t=4.20, p<0.001; self-completion>interviewer-completed** **TC: t=7.91, p<0.001; self-completion>interviewer-completed** **HC: t=10.19, p<0.001; self-completion>interviewer-completed**						**COMPARISION OF RESULTS FOR EACH GROUP BETWEEN THOSE WHO DO AND DO NOT MEET MCCB STANDARDIZATION CRITERIA** **UT: t=4.08, p<0.001; meet criteria>do not meet criteria** **TC: t=6.60, p<0.001; meet criteria>do not meet criteria** **HC: t=9.48, p<0.001; meet criteria>do not meet criteria**				
RMET-CV-R gender score	75.8% (15.1%)	81.4% (13.1%)	83.8% (11.4%)	76.2% (16.2%)	82.3% (12.8%)	86.3% (8.33%)	75.5% (14.3%)	80.5% (13.4%)	82.0% (12.8%)	75.5% (16.1%)	82.4% (13.0%)	86.6% (8.5%)	76.2% (14.0%)	80.0% (13.2%)	81.9% (12.6%)
	F=15.01, p<0.001 [HC,TC>UT]			**F=12.45, p<0.001 [HC,TC>UT]**			**F=5.17, p=0.006 [HC>UT]**			**F=15.23, p<0.001 [HC,TC>UT]**			**F=3.74, p=0.025 [HC>UT]**		
				COMPARISION OF RESULTS FOR EACH GROUP BETWEEN SELF-COMPLETION AND INTERVIEWER-COMPLETION RESPONDENTS UT: t=0.24, p=0.808; ns TC: t=0.75, p=0.453; ns **HC: t=3.01, p=0.003; self-completion>interviewer-completed**						**COMPARISION OF RESULTS FOR EACH GROUP BETWEEN THOSE WHO DO AND DO NOT MEET MCCB STANDARDIZATION CRITERIA** UT: t=-0.23, p=0.820; ns TC: t=0.94, p=0.349; ns **HC: t=3.22, p=0.002; meet criteria>do not meet criteria**				

RMET-CV-R Revised Chinese Version of the Reading the Mind in the Eyes; TC, Treated Control; UT, Untreated case; HC, Healthy Control; ns, not statistically significant
statistically significant results are **bolded**
[a] Comparisons between the 3 groups uses ANOVA; if statistically significant pairwise comparisons are made using the Tukey test

Table 4.6 Multivariate linear regression comparing RMET-CV-R scores (percent correct responses about emotions in 70 pictures) between untreated cases (UT), treated controls (TC) and healthy controls in the full sample and 4 subsamples

		All respondents who successfully completed RMET-CV-R n=472		Respondents who successfully self-completed RMET-CV-R (self-completed subsample) n=206		Respondents who successfully completed RMET-CV-R administered RMET-CV-R by interviewers (interviewer-completed subsample) n=266		Respondents who successfully completed RMET-CV-R with ≥ 5 years of schooling and <60 (educated-younger subsample) n=224		Respondents who successfully completed RMET-CV-R with <5 years of schooling or ≥60 (undereducated-older subsample) n=248	
		Coef	**P value**	**Coef**	**P value**	**Coef**	**P value**	**Coef**	**P value**	**Coef**	**P value**
Group	TC v. UT*	**0.035**	**0.039**	**0.073**	**0.006**	0.013	0.533	**0.057**	**0.020**	0.021	0.354
	HC v. UT*	**0.105**	**<0.001**	**0.136**	**<0.001**	**0.072**	**<0.001**	**0.138**	**<0.001**	**0.079**	**<0.001**
	HC v. TC*	**0.070**	**<0.001**	**0.063**	**0.003**	**0.059**	**0.004**	**0.081**	**<0.001**	**0.058**	**0.006**
Age in years		**-0.002**	**<0.001**	-0.001	0.349	**-0.003**	**<0.001**	-0.001	0.576	**-0.003**	**<0.001**
Gender (1=female*, 2=male)		**0.028**	**0.030**	-0.022	0.235	**0.061**	**<0.001**	-0.002	0.915	**0.053**	**0.004**
Years of schooling		**0.019**	**<0.001**	**0.012**	**0.006**	**0.010**	**0.004**	**0.012**	**0.009**	**0.020**	**<0.001**
Residence (1=urban*, 2=rural)		-0.040	0.081	-0.047	0.092	-0.047	0.179	-0.047	0.095	-0.060	0.125
Ethnicity (1=Han*, 2=minority group)		-0.010	0.446	0.015	0.426	-0.020	0.209	0.015	0.435	-0.029	0.079
Adjusted RMET-CV-R score from regression with 95% CI											
UT		34.9% (32.5%-37.3%)		41.9% (37.8%-46.0%)		29.9% (27.2%-32.6%)		41.1% (37.4%-44.8%)		29.3% (26.4%-32.1%)	
TC		38.4% (36.2%-40.6%)		49.2% (46.0%-52.3%)		31.2% (28.5%-34.2%)		46.8% (43.7%-49.9%)		31.4% (28.1%-34.6%)	
HC		45.4% (43.6%-47.2%)		55.5% (53.0%-57.9%)		37.1% (34.6%-39.5%)		54.9% (52.3%-57.5%)		37.2% (34.8%-39.5%)	

RMET-CV-R Revised Chinese Version of the Reading the Mind in the Eyes; TC, Treated Control; UT, Untreated case; HC, Healthy Control
statistically significant results are **bolded**
* denotes the reference group

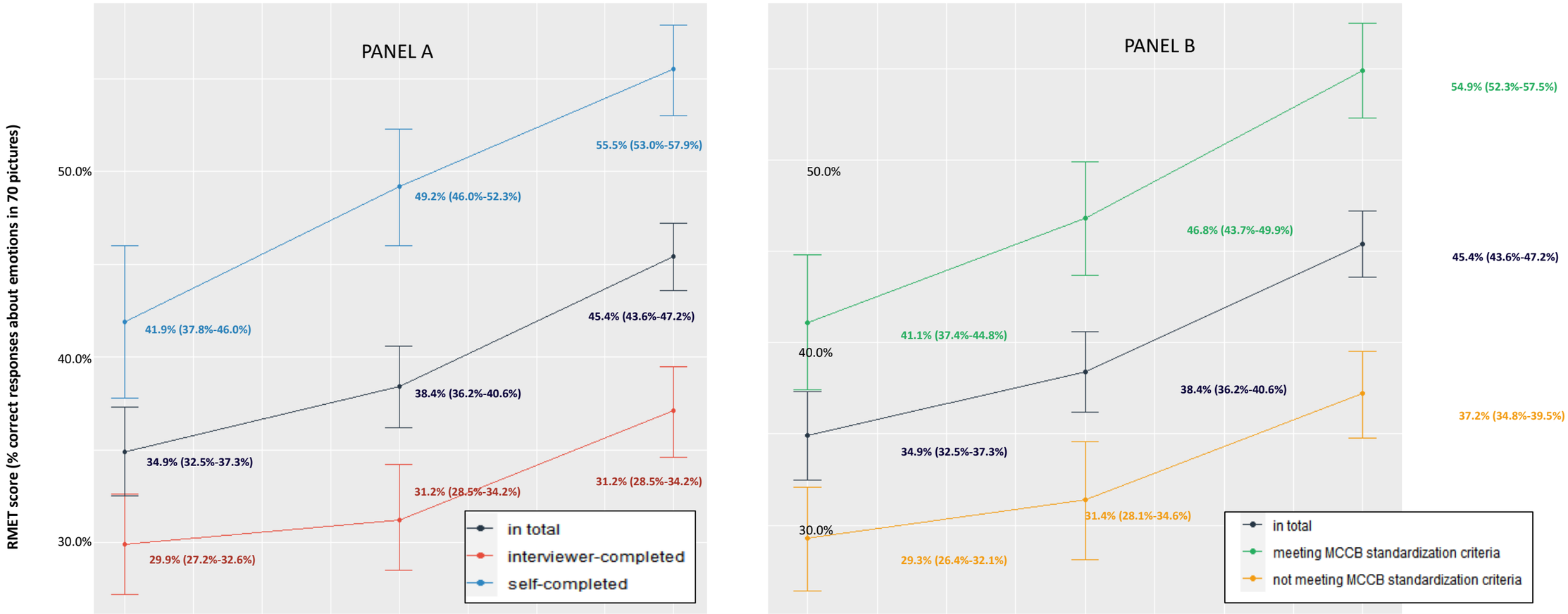

Figure 4.4 Among untreated cases, treated controls, and healthy controls who successfully complete RMET-CV-R, adjusted mean RMET-CV-R scores (percent correct responses about emotions in 70 pictures, with 95% CI) in the self-completed and interview-completed subgroups (a) and in respondents who do or do not meet MCCB standardization criteria (b) using linear models.

Heckman estimation adjusted the probability of completing the RMET test and included the participants who refused or were not able to do the test. It, as expected, revealed bigger gaps in RMET scores among the three groups (the RMET score was 8.2% higher in TC group than that in UT group, and 21.4% in HC group than that in UT group) than those using GLS model (shown in table 4.7 and figure 4.5). The correlation of the two error terms in equation five and six, ρ, was estimated to be 0.984, indicating those participants could not complete the RMET test (i.e., their RMET score was unobservable) were more likely to had lower RMET score than whose could complete the test. The Heckman model also captured the non-linear relationship between years of schooling (shown in table 4.8 and Figure 4.5), and the marginal effect of years of schooling went down with the increase of years of schooling. However, the interaction term between years of schooling and age was no longer significantly in Heckman estimation, which was possibly due to the selection bias in GLS estimation that happened more frequently in the older group was controlled by Heckman estimation. Compared with GLS estimation, the magnitude of coefficients was also larger in Heckman estimation.

Table 4.7 Multivariate linear regression comparing RMET-CV-R scores (percent correct responses about emotions in 70 pictures) between untreated cases (UT), treated controls (TC) and healthy controls in the full sample including quadratic term of years of schooling and the interaction term of age and years of schooling

		Full sample n=646	
		Coefficient	**P value**
Group (untreated group as reference group)	TC v. UT*	**0.038**	**0.024**
	HC v. UT*	**0.107**	**<0.001**
Age in years		**-0.004**	**<0.001**
Gender (1=female*, 2=male)		**0.025**	**0.049**
Years of schooling		0.013	0.225
Quadratic term of years of schooling		-0.001	0.064
Residence (1=urban*, 2=rural)		**-0.048**	**0.038**
Ethnicity (1=Han*, 2=minority group)		-0.013	0.305
Interaction: age*years of schooling		**0.0003**	**0.045**
Adjusted percent (95% CI) of successful completion of RMET-CV-R computed from regression when other covariates at mean			
UT		36.3% (33.7%-39.0%)	
TC		40.1% (37.7%-42.6%)	
HC		47.0% (44.8%-49.2%)	

RMET-CV-R Revised Chinese Version of the Reading the Mind in the Eyes; TC, Treated Control; UT, Untreated case; HC, Healthy Control

statistically significant results are **bolded**

* denotes the reference group

Table 4.8 Multivariate linear regression comparing RMET-CV-R scores (percent correct responses about emotions in 70 pictures) between untreated cases (UT), treated controls (TC) and healthy controls in the full sample including quadratic term of years of schooling and the interaction term of age and years of schooling using Heckman estimation

		Full sample n=646	
		Coefficient	**P value**
Group (untreated group as reference group)	TC v. UT*	**0.082**	**<0.001**
	HC v. UT*	**0.214**	**<0.001**
Age in years		**-0.004**	**<0.001**
Gender (1=female*, 2=male)		**0.040**	**0.004**
Years of schooling		**0.029**	**0.012**
Quadratic term of years of schooling		**-0.001**	**0.006**
Residence (1=urban*, 2=rural)		**-0.066**	**0.004**
Ethnicity (1=Han*, 2=minority group)		-0.003	0.816
Interaction: age*years of schooling		0.0002	0.267
Adjusted percent (95% CI) of successful completion of RMET-CV-R computed from regression when other covariates at mean			
UT		23.1% (20.3%-25.9%)	
TC		31.3% (28.4%-34.3%)	
HC		44.5% (42.3%-46.8%)	

RMET-CV-R Revised Chinese Version of the Reading the Mind in the Eyes; TC, Treated Control; UT, Untreated case; HC, Healthy Control

statistically significant results are **bolded**

* denotes the reference group

ρ=0.984 (95% CI 0.950-0.995)

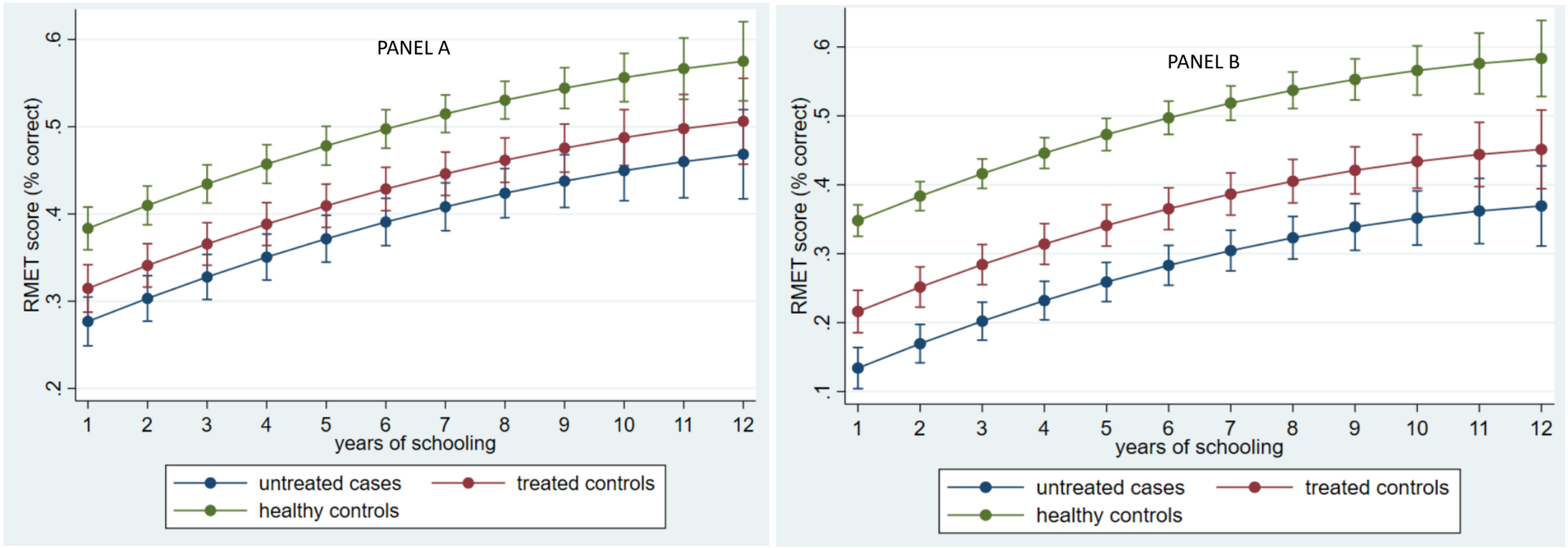

Figure 4.5 Among untreated cases, treated controls, and healthy controls who successfully complete RMET-CV-R, adjusted mean RMET-CV-R scores (percent correct responses about emotions in 70 pictures, with 95% CI) at different years of schooling using linear regression (a) and among all the untreated cases, treatment controls, and healthy controls, adjusted mean RMET-CV-R scores at different years of schooling using Heckman models.

4.4 DISCUSSION

Despite their low level of education (median of 3 years of schooling) and relatively high age (mean age=53 years), 95% of HC successfully completed RMET-CV-R, indicating the feasibility of administering RMET-CV-R to elderly, undereducated adults who are commonly excluded from studies of social cognition. After adjustment for differences in age and education in the three groups, successful completion rates in HC, TC, and UT were 96%, 69%, and 50%, respectively. The much lower successful completion rates in the patient groups suggest that failure to complete the scale successfully was *not* primarily due to low education or older age but, rather, to some aspect of the illness (or an interaction between illness factors, education, and age). In all three groups of participants, the MMSE score – a global measure of cognition – was significantly lower in those who did not successfully complete RMET-CV-R, even after adjusting for age, education, and other covariates (results provided on request). Thus, even though group differences in successful completion rates of the RMET-CV-R may not directly assess differences in social cognition, they could be considered proxies for differences in global cognitive functioning.

Given the higher proportions of UT and TC unable to successfully complete RMET-CV-R than HC, RMET-CV-R scores (based on successfully completed tests) in the patient groups exclude results for patients with the lowest levels of functioning, attenuating the difference in RMET-CV-R scores between patients and HC using GLS model. Nevertheless, patients who successfully complete RMET-CV-R have significantly lower scores than HC, confirming the findings of previous reports with RMET (Balogh et al., 2014; Caletti et al., 2013; Couture et al., 2008; Donohoe et al., 2012; Kettle et al., 2008; Mazza et al., 2013; Stanford et al., 2011; Thaler et al., 2013; Wang et al., 2006) and RMET-CV (Zhang et al., 2016; Zhang et al., 2018). This result confirms the discriminant validity of our revised version of the RMET.

In all three groups (UT, TC, and HC) the successful completion rate and the RMET-CV-R score were significantly higher among respondents who self-completed the scale than among those who had to have the interviewer read the items and significantly higher among educated-younger respondents than among undereducated-older respondents. These differences between the subsamples in the expected direction in all three groups provide further support for our revised scale's discriminant validity.

Unlike previous research that only enrolled highly educated patients or healthy control as respondents, a large percent of the patients with schizophrenia in this study were old and less educated drug-naïve patients with long duration of illness, resulting in an unbalanced completion rate of RMET test across three groups. In both the bivariate and multivariate analyses, the rate of successful completion of RMET-CV-R and the RMET-CV-R score was highest in HC, intermediate in TC, and lowest in UT in the full sample and all four subsamples. The similarity in the pattern of results between the interviewer-completed and self-completed subsamples and between the undereducated-older and educated-younger subsamples confirms the feasibility of having interviewers read RMET items to the respondent (which has not previously been attempted) and of administering the revised instrument to less-educated, older respondents (who have not previously been included in studies of RMET).

Adjusted rates of successful completion and adjusted RMET-CV-R scores were significantly higher in TC than in UT in the full sample (including all three groups) and in the analysis limited to the two patient groups with additional adjustment for duration of illness, but not significantly different when adjusting the two patient groups for both duration of illness and PANSS total score, possibly because the inclusion of some cognitive-associated symptoms in PANSS attenuates the difference in RMET-CV-R scores between UT and TC. In the subsamples, for all three analyses (i.e., the three-group analysis and two two-group analyses with additional covariates) the adjusted rate of successful completion is significantly higher in TC than in UT in the subsamples with less education (i.e., interviewer-completed and undereducated-

older subsamples), while the adjusted RMET-CV-R score is significantly higher in TC than in UT in the subsamples with more education (i.e., self-completion and educated-younger subsamples). That is, among all patient respondents, prolonged antipsychotic treatment has a greater effect on the ability to successfully complete the test in patients with lower levels of education; while among patients able to complete the test successfully, antipsychotic treatment has a greater positive effect on test scores in patients with higher levels of education. These cross-sectional findings suggest that prolonged treatment with antipsychotic medications (over 70% of the total duration of illness in this sample) could potentially mitigate deficits in social cognition in chronic schizophrenia, a possibility that needs to be confirmed in long-term prospective studies.

However, the unbalanced completion rate across three groups resulted in potential selection bias of GLS estimation about the RMET score, i.e., the adjusted RMET score in UT group was very likely to be overestimated. Therefore, Heckman estimation was used to deal with this issue. After including all the participants regardless of their completion status, the discriminant ability of the RMET scale was improved substantially and the magnitude of most coefficients became larger than that in the GLS model. The adjusted RMET score in UT turned to be much lower than that from GLS model (from 36.3% in GLS estimation to 23.1% in Heckman estimation) while the adjusted score in HC group only decreased a slight level (from 47.0% to 44.5%), reflecting the probability of participants could not complete the test due to their poor social cognition was adjusted by Heckman estimation.

Several authors have discussed concerns about the relatively strong correlation of RMET scores with vocabulary level or reading level (Olderbak et al., 2015; Pinkham et al., 2018; Eddy, 2019). In this study we read the items to those unable to read, exclude results of respondents unable to understand the descriptive terms in the slides, and adjust all results for respondents' educational level. To the extent that educational level reflects pre-illness vocabulary level, our results are also adjusted for vocabulary. The much higher proportions of UT and TC than HC unable to successfully

complete RMET-CV-R due to limited vocabulary (20%, 17%, and 3%, respectively) suggests – but does not prove – that patients' vocabulary level declined over the course of their illness; that is, their current impoverished vocabulary (compared to healthy controls with similar levels of education) was not part of their pre-illness status. If one supposes that lack of related vocabulary is a marker of inability to name emotions (and, thus, to attribute the mental state of another), making further adjustments of RMET-CV-R scores for vocabulary level could result in over-control, obscuring real differences in social cognition between groups.

Another intriguing finding is the role of education and age. As shown in Tables 4.3 and 4.5, in the full sample, education and age are significantly associated with successful completion of RMET-CV-R and RMET-CV-R scores; however, the pattern changes when considering the subsamples. Age and level of education were more closely associated with successful completion and RMET-CV-R scores in the interviewer-completed and undereducated-older subsamples than in the self-completed and educated-younger subsamples. This suggests that among subjects typically included in previous studies (i.e., respondents <60 years of age with $\geq$5 years of schooling who can self-complete the instrument), older age and lower levels of education have relatively little independent effect on rates of successful completion and RMET-CV-R scores. However, when considering older, less-educated respondents and those unable to self-complete the instrument, increasing age and lower levels of schooling are independently associated with greater deficits in cognitive functioning (i.e., lower rates of successful completion and lower RMET-CV-R scores) even after adjusting for group (UT, TC, and NC) and other covariates (Staekenborg et al., 2020).

4.5 Limitations

There are several potential limitations. 1) We have not yet assessed the test-retest reliability or other psychometric characteristics of RMET-CV-R. 2) The enrolled UT and TC may not be representative of UT and TC in other locations. 3) Our relatively large sample made it feasible to adjust results for intergroup differences in

demographic and clinical characteristics, but there may still be unassessed differences between the groups. For instance, the group assignment between treated control and untreated patients were not random, so there might be other unobservable factors related to both group assignment and the outcome variable. 4) This is a cross-sectional study so no definitive causal conclusions can be made about the factors associated with successful completion of RMET-CV-R or RMET-CV-R scores in the multivariate analyses.

4.6 Conclusion

Social cognition is an important target for therapeutic interventions such as social skills training and cognitive remediation. Assessing these interventions' effectiveness requires instruments that can validly assess social cognition in all patients with schizophrenia, particularly less-educated, older patients who tend to have more severe deficits in social cognition. Despite the limitations of the current study, we conclude that with appropriate changes in the method of test administration and careful assessment of respondents' understanding and ability to complete the task required, it is possible to administer RMET to many individuals with schizophrenia who were previously considered ineligible. This revised method of administering and evaluating RMET can potentially be used to assess the effectiveness of cognitive remediation interventions among less-educated, older patients.

This result has important implications for all research about schizophrenia. Given the increasing proportion of individuals with schizophrenia in high-income countries who are older (approaching the age where the onset of dementia needs to be considered) and the large proportion of individuals with schizophrenia in LMIC with low levels of education, it is no longer appropriate to automatically exclude them from research studies about schizophrenia due to their age or level of education. The symptomatology, course, and treatment responsiveness of older, undereducated

patients – particularly those from rural communities – may differ from those of younger, educated patients (Yang et al., 2020). To ensure that the needs of these marginalized patients are addressed, researchers need to make the added effort to include them in their studies. A critical part of this effort is to adapt assessment instruments to make them suitable for use with undereducated, older patients. We believe that this study represents an important step towards this goal.

Acknowledgements

The authors thank the psychiatric clinicians who assessed and verified the DSM-IV diagnoses using SCID-IV for this study, including Drs. Jizu Wei and Guoqiang Lü from Ningxia Guyuan City Yuanzhou District Psychiatric Rehabilitation Hospital, Drs. Xingbin Duan, Yanhua Chen, Yujian Wu and Lijuan Xi from Ningxia Minkang Hospital, Drs. Weiguo Xu, Xuelan Liu, Huanqiu Wang from Ningxia Ningan Hospital, and Dr. Zhaokui Yang from Guangxi Brain Hospital. We thank staff members at the local Centers of Disease Control and other health workers around Ningxia and Guangxi who helped coordinate the fieldwork. We thank the patients, patients' guardians, and healthy control subjects who agreed to participate in the study. Finally, we gratefully acknowledge Dr. Sarah Morris at the National Institutes of Mental Health for her support.

The two senior authors for the paper, Dr. Yang (New York University and Columbia University) and Dr. Phillips (Shanghai Mental Health Center and Columbia University), had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Financial support

This study was supported by the National Institute of Mental Health grant MH108385 R01 (PI's: Yang, Phillips, Keshavan) for the project 'Characterizing

Cognition Across the Lifespan in Untreated Psychosis in China'. The NIMH had no other role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Conflicts of Interest

The authors declare no potential conflicts of interest with respect to the research, authorship, and publication of this article.

Ethical standards

This study was approved by the institutional review boards of Shanghai Mental Health Center (FWA 00003065, IOGA 0002202) and New York University (IRB-FY2016-1157). All participants provided written informed consent; in the two patient groups, guardians also signed written informed consent forms.

REFERENCES

Achim, A.M., Ouellet, R., Roy, M.A., Jackson, P.L., 2012. Mentalizing in first-episode psychosis. *Psychiatry Research*. 196, 207–213. doi: http://dx.doi.org/10.1016/j.psychres. 2011.10.011

Balogh, N., Égerházi, A., Berecz, R., & Csukly, G. (2014). Investigating the state-like and trait-like characters of social cognition in schizophrenia: A short term follow-up study. *Schizophrenia Research*, *159*(2-3), 499-505. doi: https://doi.org/10.1016/j.schres.2014.08.027

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The Reading the Mind in the Eyes Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *42*(2), 241-251. doi: https://doi.org/10.1111/1469-7610.00715

Bell, M. D., Corbera, S., Johannesen, J. K., Fiszdon, J. M., & Wexler, B. E. (2013). Social cognitive impairments and negative symptoms in schizophrenia: are there subtypes with distinct functional correlates?. *Schizophrenia Bulletin*, *39*(1), 186-196. doi: https://doi.org/10.1093/schbul/sbr125

Bora, E., Yucel, M., & Pantelis, C. (2009). Theory of mind impairment in schizophrenia: meta-analysis. *Schizophrenia research*, *109*(1-3), 1-9. doi: https://doi.org/10.1016/j.schres.2008.12.020

Caletti, E., Paoli, R. A., Fiorentini, A., Cigliobianco, M., Zugno, E., Serati, M., ... & Altamura C. A. (2013). Neuropsychology, social cognition and global functioning among bipolar, schizophrenic patients and healthy controls: preliminary data. *Frontiers in human neuroscience*, 7, 1-14. doi: 10.3389/fnhum.2013.00661

Carter, C. S., Barch, D.M., Gur, G., Gur, R., Pinkham, A., & Ochsner, K. (2009) CNTRICS final task selection: social cognitive and affective neuroscience-based measures. *Schizophrenia Bulletin, 35*(1), 153-162. doi:

https://doi.org/10.1093/schibul/sbn152

Corcoran, R., Mercer, G., & Frith, C. D. (1995). Schizophrenia, symptomatology and social inference: investigating “theory of mind” in people with schizophrenia. *Schizophrenia Research*, *17*(1), 5-13. doi: https://doi.org/10.1016/0920-9964(95)00024-G

Couture, S. M., Penn, D. L., Addington, J., Woods, S. W., & Perkins, D. O. (2008). Assessment of social judgments and complex mental states in the early phases of psychosis. *Schizophrenia Research*, *100*(1-3), 237-241. doi: https://doi.org/10.1016/j.schres.2007.12.484

Dinno, A. (2015) Nonparametric pairwise multiple comparisons in independent groups using Dunn’s test. The Stata Journal, 15(1): 292-300

Donohoe, G., Duignan, A., Hargreaves, A., Morris, D. W., Rose, E., Robertson, D., ... & Corvin, A. (2012). Social cognition in bipolar disorder versus schizophrenia: comparability in mental state decoding deficits. *Bipolar disorders*, *14*(7), 743-748. doi: https://doi.org/10.1111/bdi.12011

Dunn, O.J. (1961) Multiple comparisons among means. Journal of the American Statistical Association 56: 52-64

Eddy, C.M. (2019) What do you have in mind? Measures to assess mental state reasoning in neuropsychiatric populations. *Frontiers in Psychiatry, 10: 425.* doi: https://doi.org/10.3389/fpsyt.2019.00452

Fett, A.K.J., Viechtbauer, W., Dominguez, M., Penn, D. L., van Os, J., & Krabbendam, L. (2011). The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: a meta-analysis. *Neuroscience & Biobehavioral Reviews*, *35*(3), 573-588. doi: https://doi.org/10.1016/j.neubiorev.2010.07.001

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”: a practical method for grading the cognitive state of patients for the clinician. *Journal of psychiatric research*, *12*(3), 189-198. doi: https://doi.org/10.1016/0022-

3956(75)90026-6

Green, M.F., Horan, W.P., & Lee, J. (2015) Social cognition in schizophrenia. *Nature Reviews Neuroscience,* 16(10), 620-631. doi: https://doi.org/10.1038/nrn4005

Green, M.F., Horan, W.P., & Lee, J. (2019) Nonsocial and social cognition in schizophrenia: current evidence and future directions. *World Psychiatry,* 18, 146-161. doi: https://doi.org/10.1002/wps.20624

Guastella, A. J., Hermens, D. F., Van Zwieten, A., Naismith, S. L., Lee, R. S. C., Cacciotti-Saija, C., ... & Hickie, I. B. (2013). Social cognitive performance as a marker of positive psychotic symptoms in young people seeking help for mental health problems. *Schizophrenia research*, *149*(1-3), 77-82. doi: https://doi.org/10.1016/j.schres.2013.06.006

Harvey, P. D., & Rosenthal, J. B. (2018) Cognitive and functional deficits in people with schizophrenia: Evidence for accelerated or exaggerated aging? *Schizophrenia Research,* 196, 14-21. doi: https://doi.org/10.1016/j.schres.2017.05.009

Harvey, P. D., Strassnig, M. T., & Silberstein, J. (2019). Prediction of disability in schizophrenia: Symptoms, cognition, and self-assessment. *Journal of Experimental Psychopathology*, *10*(3), 1-20. doi: https://doi.org/10.1177/2043808719865693

Kay, S. R., Fiszbein, A., & Opler, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia bulletin*, *13*(2), 261-276. doi: https://doi.org/10.1093/schbul/13.2.261

Kettle, J. W., O'Brien-Simpson, L., & Allen, N. B. (2008). Impaired theory of mind in first-episode schizophrenia: comparison with community, university and depressed controls. *Schizophrenia research*, *99*(1-3), 96-102. doi: https://doi.org/10.1016/j.schres.2007/11.011

Mayer, J. D., Salovey, P., Caruso, D. R., & Sitarenios, G. (2003). Measuring emotional intelligence with the MSCEIT V2. 0. *Emotion*, *3*(1), 97-105. doi: http://doi.org/10.1037/1528-3542.3.1.97

Mazza, M., Catalucci, A., Pino, M. C., Giusti, L., Nigri, A., Pollice, R., ... & Gallucci, M. (2013). Dysfunctional neural networks associated with impaired social interactions in early psychosis: an ICA analysis. *Brain imaging and behavior*, *7*(3), 248-259. doi: https://doi.org/10.1007/s11682-013-9223-6

Mesholam-Gately, R. I., Giuliano, A. J., Goff, K. P., Faraone, S. V., & Seidman, L. J. (2009). Neurocognition in first-episode schizophrenia: a meta-analytic review. *Neuropsychology*, *23*(3), 315-336. doi: https://doi.org/10.1037/a0014708

Olderbak, S., Wilhelm, O., Olaru, G., Geiger, M., Brenneman, M.W., & Roberts, R.D. (2015) A psychometric analysis of the reading the mind in the eyes test: towards a brief form for research and applied settings. *Frontiers in Psychology, 6:* 1503. doi: https://doi.org/10.3389/fpsyg.2015.01503

Phillips MR, Liu XH. Translated and adapted Chinese version of Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition (SCID-I/P) by Michael B. First, Robert L. Spitzer, Miriam Gibbon, and Janet B.W. Williams. Shanghai: Suicide Research and Prevention Center, Shanghai Mental Health Center; 2011.

Pinkham, A. E., Harvey, D. H., & Penn, D. L. (2018) Social cognition psychometric evaluation: results of the final validation study. *Schizophrenia Bulletin, 44*(4), 737-748. doi: https://doi.org/10.1093/schibul/sbx117

Pinkham, A. E., Penn, D.L., Green, M.F., & Harvey, P.D. (2016) Social cognition psychometric evaluation: results of the initial psychometric study. *Schizophrenia Bulletin, 42*(2), 494-504. doi: https://doi.org/10.1093/schibul/sbv056

Sheffield, J. M., Karcher, N. R., & Barch, D., M. (2018) Cognitive deficits in psychotic disorders: a lifespan perspective. *Neuropsychology Review*, 28, 509-533. doi: https://doi.org/10.1007/s11065-018-9388-2

Shi, C., Kang, L., Yao, S., Ma, Y., Li, T., Liang, Y., ... & Yu, X.. (2015). The MATRICS Consensus Cognitive Battery (MCCB): co-norming and standardization in China. *Schizophrenia research*, *169*(1-3), 109-115. doi:

https://doi.org/10.1016/j.schres.2015.09.003

Staekenborg, S. S., Kelly, N., Schuur, J., Koster, P., Scherder, E., Tielkes, C. E. M…….& Claus, J.J. (2020) Education as a proxy for cognitive reserve in a large elderly memory clinic: window of benefit. *Journal of Alzheimer's Disease*, 2020, 76(2), 671-679. doi: https://doi.org/10.3233/JAD-191332

Stanford, A. D., Messinger, J., Malaspina, D., & Corcoran, C. M. (2011). Theory of Mind in patients at clinical high risk for psychosis. *Schizophrenia research*, *131*(1-3), 11-17. doi: https://doi.org/10.1016/j.schres.2011.06.005

Stone, W. S., Cai, B., Liu, X., Grivel, M. M. R., Yu, G., Xu, Y., ... & Phillips, M.R. (2020). Association Between the Duration of Untreated Psychosis and Selective Cognitive Performance in Community-Dwelling Individuals With Chronic Untreated Schizophrenia in Rural China. *JAMA Psychiatry.* Advance online publication. doi: https://doi.org/10.1001/jamapsychiatry.2020.1619

Thaler, N. S., Allen, D. N., Sutton, G. P., Vertinski, M., & Ringdahl, E. N. (2013). Differential impairment of social cognition factors in bipolar disorder with and without psychotic features and schizophrenia. *Journal of Psychiatric Research*, *47*(12), 2004-2010. doi: https://doi.org/10.1016/j.jpsychires.2013.09.010

Vohs, J. L., Lysaker, P. H., Francis, M. M., Hamm, J., Buck, K. D., Olesek, K., ... & Breier, A. (2014). Metacognition, social cognition, and symptoms in patients with first episode and prolonged psychoses. *Schizophrenia Research*, *153*(1-3), 54-59. doi: https://doi.org/10.1016/j.schres.2014.01.012

Wang, Y. G., Wang, K., Zhu, C. Y., & Wang, H. (2006). A study of theory of mind deficit in first-episode schizophrenic patients (in Chinese). *Chinese Journal of Nervous and Mental Diseases*, *32*(6), 481-486. Retrieved from http://www.cnki.com.cn/Article/CJFD_SIMINDEX_E-ZSJJ200606002.htm

Wu, X.M., Ma, N., Wang, X., Zhang, W.F., Guan, L,L,, Ma, H., ….Lu, L. (2019) Management and service for psychosis in People's Republic of China in 2017 (in Chinese). *Chinese Journal of Psychiatry*, 52(1), 82-88. doi:

https://doi.org/10.3760/cma.j.issn.1006-7884.2019.01.008

Yang, L., Ruiz, B., Mandavia, A.D., Grivel, M.M., Wong, L.Y., Phillips, M., & … Stone, W. (2020) Advancing study of cognitive impairments for antipsychotic-naïve psychosis in low- and middle-income countries: systematic review of cognition and study methodology. *Schizophrenia Research*, 220, 1-15. doi: http://dx.doi.org/10.1016/j.schres.2020.01.026.

Zar, J. H. (1999). *Biostatistical Analysis*. Upper Saddle River, N.J.: Prentice-Hall.

Zhang, T.H., Tang, Y.Y., Cui, H.R., Lu, X., Hu, L.H., Liu, X.H.,…Wang, J.J. (2016) Theory of mind impairments in youth at clinical high risk of psychosis. *Psychiatry*, 79(1), 40-55. doi: https://doi.org/10.1016/j.psychres.2018.02.008

Zhang, T.H., Cui, H., Wei, Y., Tang, Y., Xu, L., Tang, X., ... Wang, J.J. (2018). Progressive decline of cognition during the conversion from prodrome to psychosis with a characteristic pattern of the theory of mind compensated by neurocognition. *Schizophrenia Research*, *195*, 554-559. doi: https://doi.org/10.1080/00332747.2015.

SUPPLEMENTARY MATERIALS of Chapter Four

Changes in the administration of the RMET-CV-R

The original 36-item RMET and the 70-item Chinese version (RMET-CV) are computerized tests which are self-completed by respondents. However, in this study 59.4% of the sample (n=384) were over 60 years of age or had less than 5 years of schooling, and 88% had never used a computer. To conduct the test with these respondents it was necessary to make the following changes in the administration of the test (without changing the presented pictures or optional descriptors provided in the original versions of the test).

1) All RMET-CV-R tests are conducted in the presence of an interviewer. Unlike the administration of the RMET, for the RMET-CV-R the interviewer reads aloud the requirements of the test and shows the respondent how to make responses during the training session prior to starting the main test.

2) For all participants with less than five years of schooling, the two training items and the 70 items in the RMET-CV-R main test were 'interviewer-completed': the interviewer reads aloud the four options for each picture at a standard speed (1 second for two-Chinese-character-words, 2 seconds for four-Chinese-character-words) and clicked the mouse to select the descriptor chosen by the respondent. The participant was then asked if the picture was of a male or a female and the interviewer clicked the corresponding response on the screen.

3) For participants with five or more years of schooling, they are first trained to use the computer mouse by having them independently manipulate the mouse to move the position of the cursor over 5 numbered buttons on the computer screen and click each button in sequence. If the respondent was unable to complete this task after several trials, the test was interviewer-completed.

4) Presentation of the two pictures used for training in the RMET-CV-R is

managed by the interviewer who explains the expectations at each step (i.e., to select an appropriate emotional descriptor and then indicate the gender of the presented eyes) and asks the respondent to make a selection when presented with the pictures. If the respondent has difficulty in completing this task, the interviewer makes up to 6 attempts to help the respondent understand the requirements of the task prior to starting the main trial.

5) Respondents with more than five years of education who are able to learn how to use a mouse are expected to complete the computerized main test on their own (i.e., a 'self-completed' test). However, if a respondent clearly has difficulty reading the words on the slide or manipulating the mouse either during the training slides or during the main test, the interviewer reverts to the 'interviewer-completed' approach.

6) One of the main difficulties these respondents had completing the test was that they often did not understand some of the emotional descriptors used. In the RMET and some of the other translated versions of the instrument (such as the Spanish and Italian version provided at https://www.autismresearchcentre.com/tests/eyes-test-adult/), respondents are provided a definition handout for terms used in the slides that they can refer to while doing the test; but this was not done in the RMET-CV or other translated versions (such as the German and French versions, provided at https://www.autismresearchcentre.com/tests/eyes-test-adult/). Similarly, in our revised Chinese version (RMET-CV-R) we have not provided a definition handbook because many of the respondents were illiterate or semi-literate, so having the interviewer read the definitions would have inordinately extended the duration of the test. However, to reduce the number of respondents who were excluded because of limited vocabulary, respondents are allowed to skip up to 4 of the 70 pictures (skipped items are considered 'incorrect' when computing total number of correct responses). When the respondent fails to complete more than 4 items, the test is terminated and considered

incomplete.

7) To balance any potential order effects, 327 participants enrolled before September 2017 were first presented Asian then Caucasian pictures, while 320 participants enrolled after September 2017 were first presented Caucasian then Asian pictures.

8) Participants who have started the main test but refuse to continue before finishing are encouraged a maximum to two times with the phrase 'please try to complete the test'; if they persist in their request to stop the test, the test is terminated and the results considered incomplete.

The computerized RMET-CV-R was generated using Java. After completion or termination of each test a separate text file recording the selected emotional descriptor, the selected gender, and the time to make both of these decisions for each of the 70 pictures and a summary table separating results for the 36 Caucasian and 34 Asian pictures is downloaded. These data were then automatically extracted into a csv database using Python.

Assessment of RMET-CV-R completion status

Given the possibility that some older or uneducated respondents could not successfully complete RMET-CV-R, we classified the completion status of each administration of the test based on the respondent's willingness to complete the test, ability to complete the test, and understanding of the required task, as follows:

a) Respondent explicitly refuses to complete the practice examples, to start the main test *or* to respond to 1 or more pictures during a main test in which a total of 5 pictures have been skipped (either due to refusal or failure to select a response).

b) Respondent skips a total of 5 pictures (occurring at any point in the series of 70 pictures) none of which are skipped due to active refusal *and* has indicated in some way that he/she understands what is expected (either by choosing an

emotional descriptor or by stating that they are unable to choose a descriptor for any slide). In almost all cases this occurs because the respondent has limited vocabulary and, thus, does not understand the provided emotional descriptors.

c) Respondent provides no active response during the practice examples or during the first five pictures in the main test *and* has not indicated in any way that he/she at least partially understands what is expected.

d) Respondent completes 66-70 pictures.

Tests classified as ‘a’, ‘b’ or ‘c’ are considered ‘incomplete’ and those classified as ‘d’ are considered ‘successfully completed’.

Supplementary Table 4.1A Multivariate logistic regression results comparing rates of successful completion of RMET-CV-R between untreated cases (UT) and treated controls (TC) in the full sample and in different subsamples, including duration of illness as a covariate

	Total sample[a] n=387		Respondents assigned to self-completed RMET-CV-R (self-completion subsample) n=136		Respondents assigned to have RMET-CV-R administered by interviewer (interviewer-completed subsample) n=251		Respondents with ≥ 5 years of schooling and <60 (educated-younger subsample) n=163		Respondents with <5 years of schooling or ≥ 60 (undereducated-older subsample) n=224	
	OR	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**
Group TC v. UT*	**2.58**	**1.58-4.22**	1.63	0.68-3.91	**3.18**	**1.75-5.76**	1.91	0.80-3.72	**3.12**	**1.64-5.94**
Age in years	0.99	0.96-1.03	1.00	0.92-1.09	0.99	0.96-1.02	0.98	0.92-1.04	0.99	0.96-1.03
Gender (1=female*, 2=male)	**1.77**	**1.08-2.91**	2.27	0.96-5.39	1.58	0.85-2.94	1.60	0.75-3.39	1.83	0.91-3.69
Years of schooling	1.23	1.14-1.34	1.29	1.01-1.63	**1.28**	**1.11-1.46**	1.23	1.00-1.52	1.40	1.16-1.69
Residence (1=urban*, 2=rural)	0.41	0.15-1.13	0.82	0.19-3.48	**0.26**	**0.08-0.87**	0.55	0.11-2.83	0.34	0.09-1.21
Ethnicity (1=Han*, 2=minority group)	1.22	0.75-1.99	1.06	0.39-2.90	1.27	0.72-2.26	1.41	0.57-3.51	1.15	0.62-2.13
Duration of illness (in years)	**0.96**	**0.94-0.99**	0.96	0.90-1.03	**0.96**	**0.93-0.99**	0.99	0.93-1.05	**0.95**	**0.92-0.98**
Adjusted percent (95% CI) of successful completion of RMET-CV-R computed from regression										
UT	50.9% (45.2%-56.6%)		73.8% (64.2%-83.3%)		38.7% (31.8%-45.6%)		71.8% (62.7%-80.9%)		35.8% (28.6%-43.0%)	
TC	68.9% (62.0%-75.9%)		81.4% (71.9%-90.9%)		62.5% (52.7%-72.2%)		82.4% (74.1%-90.8%)		59.2% (48.9%-69.5%)	

RMET-CV-R Revised Chinese Version of the Reading the Mind in the Eyes

statistically significant results are **bolded**

* denotes the reference group

Supplementary Table 4.1B Multivariate logistic regression results comparing rates of successful completion of RMET-CV-R between untreated cases (UT) and treated controls (TC) in the full sample and in different subsamples, including duration of illness and PANSS total score as covariates

	Full sample n=387		Respondents assigned to self-completed RMET-CV-R (self-completion subsample) n=136		Respondents assigned to have RMET-CV-R administered by interviewer (interviewer-completed subsample) n=251		Respondents with ≥ 5 years of schooling and <60 (educated-younger subsample) n=163		Respondents with <5 years of schooling or ≥ 60 (undereducated-older subsample) n=224	
	OR	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**
Group TC v. UT*	1.57	0.91-2.71	0.57	0.20-1.62	**2.16**	**1.15-4.05**	0.63	0.24-1.61	**2.28**	**1.18-4.42**
Age in years	1.00	0.96-1.03	0.99	0.90-1.10	0.99	0.96-1.02	1.00	0.93-1.08	0.99	0.96-1.03
Gender (1=female*, 2=male)	1.58	0.94-2.66	2.59	0.98-6.85	1.37	0.72-2.61	1.84	0.79-4.29	1.61	0.80-3.27
Years of schooling	**1.22**	**1.12-1.33**	**1.36**	**1.06-1.76**	**1.28**	**1.12-1.45**	**1.33**	**1.05-1.67**	**1.39**	**1.15-1.68**
Residence (1=urban*, 2=rural)	0.39	0.14-1.08	0.58	0.15-2.26	**0.26**	**0.08-0.87**	0.29	0.07-1.20	0.34	0.09-1.24
Ethnicity (1=Han*, 2=minority group)	1.30	0.76-2.22	0.65	0.22-1.97	1.44	0.78-2.67	0.79	0.29-2.10	1.31	0.68-2.51
Duration of illness	**0.97**	**0.96-1.00**	0.97	0.90-1.04	0.97	0.93-1.00	0.98	0.93-1.04	**0.95**	**0.92-0.99**
PANSS total score	**0.98**	**0.96-0.98**	**0.94**	**0.91-0.96**	**0.97**	**0.96-0.99**	**0.93**	**0.91-0.96**	**0.98**	**0.96-0.99**
Adjusted percent (95% CI) of successful completion of RMET-CV-R computed from regression										
UT	55.0% (49.6%-60.4%)		79.6% (72.1%-87.2%)		41.9% (34.9%-49.0%)		78.9% (71.9%-85.9%)		38.4% (30.9%-45.8%)	
TC	62.8% (55.4%-70.3%)		72.5% (62.5%-82.6%)		56.9% (47.2%-66.6%)		72.7% (63.2%-82.3%)		54.6% (44.3%-64.8%)	

RMET-CV-R Revised Chinese Version of the Reading the Mind in the Eyes; PANSS, Positive and Negative Syndrome Scale

statistically significant results are **bolded**

* denotes the reference group

Supplementary Table 4.2A. Multivariate linear regression comparing RMET-CV-R scores (percent correct responses about emotions in 70 pictures) between untreated cases (UT) and treated controls (TC) in the full sample and in different subsamples, including duration of illness as a covariate

	All respondents who successfully completed RMET-CV-R n=224		Respondents who successfully self-completed RMET-CV-R (self-completion subsample) n=105		Respondents who successfully completed RMET-CV-R administered RMET-CV-R by interviewers (interviewer-completion subsample) n=119		Respondents who successfully completed RMET-CV-R who were with ≥ 5 years of schooling and <60 (educated-younger subsample) n=125		Respondents who successfully completed RMET-CV-R with <5 years of schooling or ≥ 60 (undereducated-older subsample) n=99	
	Coef	**P value**	**Coef**	**P value**	**Coef**	**P value**	**Coef**	**P value**	**Coef**	**P value**
Group TC v. UT*	**0.037**	**0.031**	**0.069**	**0.010**	0.012	0.598	**0.057**	**0.022**	0.017	0.464
Age in years	**-0.002**	**0.043**	-0.0001	0.977	-0.002	0.096	-0.0005	0.807	-0.003	0.052
Gender (1=female*, 2=male)	**0.041**	**0.031**	0.018	0.518	**0.053**	**0.044**	0.033	0.202	0.044	0.134
Years of schooling	**0.015**	**<0.001**	0.010	0.105	0.010	0.092	0.008	0.232	**0.016**	**0.009**
Residence (1=urban*, 2=rural)	**-0.073**	**0.004**	**-0.085**	**0.011**	-0.054	0.149	**-0.081**	**0.022**	**-0.082**	**0.039**
Ethnicity (1=Han*, 2=minority group)	-0.024	0.156	-0.007	0.817	-0.033	0.118	-0.003	0.896	**-0.051**	**0.031**
Duration of illness (in years)	-0.002	0.816	-0.0005	0.829	-0.0002	0.848	-0.001	0.533	-0.00001	0.965
Adjusted mean (95% CI) RMET-CV-R score computed from regression										
UT	36.6% (34.1%-39.1%)		42.6% (38.4%-46.8%)		31.1% (28.2%-33.9%)		41.4% (37.6%-45.1%)		30.2% (27.3%-33.1%)	
TC	40.3% (38.1%-42.5%)		49.5% (46.4%-52.6%)		32.2% (29.1%-35.3%)		47.1% (44.0%-50.1%)		31.9% (28.5%-35.2%)	

RMET-CV-R Revised Chinese Version of the Reading the Mind in the Eyes
statistically significant results are **bolded**
* denotes the reference group

Supplementary Table 4.2B Multivariate linear regression comparing RMET-CV-R scores (percent correct responses about emotions in 70 pictures) between untreated cases (UT) and treated controls (TC) in the full sample and in different subsamples, including duration of illness and PANSS total score as covariates

	All respondents who successfully completed RMET-CV-R n=224		Respondents who successfully self-completed RMET-CV-R (self-completion subsample) n=105		Respondents who successfully completed RMET-CV-R administered RMET-CV-R by interviewers (interviewer-completion subsample) n=119		Respondents who successfully completed RMET-CV-R who were with ≥ 5 years of schooling and <60 (educated-younger subsample) n=125		Respondents who successfully completed RMET-CV-R with <5 years of schooling or ≥ 60 (undereducated-older subsample) n=99	
	Coef	**P value**	**Coef**	**P value**	**Coef**	**P value**	**Coef**	**P value**	**Coef**	**P value**
Group TC v. UT*	**0.025**	**0.163**	**0.067**	**0.016**	-0.001	0.974	**0.051**	**0.045**	0.004	0.865
Age in years	-0.002	0.059	0.001	0.963	-0.002	0.084	-0.0003	0.889	**-0.003**	**0.045**
Gender (1=female*, 2=male)	**0.040**	**0.033**	0.018	0.519	**0.052**	**0.043**	0.033	0.204	0.045	0.118
Years of schooling	**0.014**	**<0.001**	0.010	0.109	0.006	0.111	0.008	0.217	**0.015**	**0.018**
Residence (1=urban*, 2=rural)	**-0.075**	**0.003**	**-0.085**	**0.012**	-0.059	0.113	**-0.082**	**0.020**	**-0.084**	**0.034**
Ethnicity (1=Han*, 2=minority group)	-0.023	0.161	-0.007	0.811	-0.031	0.141	-0.005	0.859	**-0.047**	**0.045**
Duration of illness (in years)	-0.0002	0.823	-0.0005	0.828	-0.002	0.850	-0.001	0.535	-0.0003	0.977
PANSS total score	-0.001	0.080	-0.0001	0.908	-0.001	0.091	-0.0004	0.632	-0.001	0.096
Adjusted mean (95% CI) RMET-CV-R score computed from regression										
UT	37.2% (34.7%-39.7%)		42.7% (38.5%-46.9%)		31.7% (28.7%-34.7%)		41.7% (37.9%-45.4%)		30.8% (27.7%-33.9%)	
TC	39.6% (37.4%-41.9%)		49.4% (46.1%-52.7%)		31.6% (28.6%-34.6%)		46.8% (43.6%-49.9%)		31.2% (27.9%-34.6%)	

RMET-CV-R Revised Chinese Version of the Reading the Mind in the Eyes; PANSS, Positive and Negative Syndrome Scale
statistically significant results are **bolded**
* denotes the reference group

Chapter Five: Conclusion

This whole chapter depicts the comprehensive living status of patients with chronic schizophrenia living the community of a particular areas in rural China, including their service utilization, medication history, interaction with community health system, symptom, cognition, and so forth, which is very valuable when most research focus on the clinical outcomes and limit to the patients in the hospitals.

Chapter two conducted a large-scale empirical study (n=2280, covering all the patients with schizophrenia in eight counties) and detailed analysis about the implementation of the national community mental health program in a province in northwest China.

First of all, this study showed great improvement in community mental health provision and utilization in rural China. After over 10 years of development, the community mental health program, the National Information System of Patients with Psychosis, has covered more than 99% of the counties in China. In Ningxia, where this study was conducted, NISPP has covered all the townships or urban communities since 2010 to help patients with severe mental illness to obtain formal diagnosis and free antipsychotics in the community without going to the mental hospitals. The treatment gap in patients with schizophrenia was only 6.5% when the study was conducted and about two-thirds of the patients was regularly using antipsychotic, which reflected the mental health service utilization in rural China had made significant progress.

Secondly, this study also confirmed that the inequity of mental health service utilization still existed. Patients living in the places with longer distance to the nearest psychiatric or living in the country without mental hospital in the county had lower probability of ever initiating the treatment and having good compliance of drug use. Besides the environmental factors, personal factors also influenced the mental health service utilization. Older age, less years of formal education, being ethic minority group were all related to lower probability of ever starting the treatment and keep good compliance. Furthermore, we found ‘substitution effect’ between personal

factors and environmental factors. In the areas with better environmental factors, the majority of patients would have access to mental health service regardless their personal factors, but in the places with worse environmental characteristics, patients had to be with better personal factor to access the service.

Furthermore, we found that community mental health program, to some extent, violated its initial goal to improve the equity across different areas. For instance, more free drugs were provided to urban areas closer to mental hospitals. The potential reason is in urban area, there are more human resource to utilize the budget of community program and provide this community service. The original inequity to eliminate the inequity was violated due to the unequal distribution of human resource and enlarged the gap between urban and rural areas by implementing the public health program. This dilemma might not only happen when implementing community mental health program in rural China, and it would widely exist in all developing countries.

Although this study confirmed the positive effect of implementing community mental health program in rural China, there is still no rigorous evaluation measuring the effectiveness of NISPP which is the largest community mental health program in the world. One key obstacle for evaluation is the methodologically issue. Different from traditional clinical trials, it is hard to find matched control groups for policy and service interventions and apply standard methods, such as randomized controlled trials, so it is suitable to use innovative econometric models in the future research.

Chapter Three is a systematic review to compare a component of social cognition, theory of mind, measured by Reading the Mind in the Eyes Scale (RMET) in patients with schizophrenia and healthy controls, and also to explore the effect of age and years of schooling on RMET score separately in two groups of participants, patient with schizophrenia and healthy controls.

In total, 198 studies met the inclusion criteria including 41 separate samples of patients with schizophrenia (n=1836) and 197 separate samples of healthy controls (n=23,675). As expected, the 26 studies directly compared RMET results in patients and healthy controls indicated that RMET scores in patients were 1.10 mean standard

deviations lower than RMET scores in healthy controls ($p<0.001$), indicating patients with schizophrenia showed substantial deficits compared with healthy control. Meta regression separately in patients and healthy controls, found a negative association of age with RMET score and a positive association of years of schooling with RMET score. A secondary meta-analysis using a spline construction of the 180 healthy control samples that include data on mean age identified a non-monotonic relationship between age and RMET score – RMET scores increased with age before age 31 and decreased with age after age 31, which provides a potential new approach for assessing how theory of mind changes with aging.

This research fits in an important argument about schizophrenia, whether schizophrenia is a neurodegenerative or neurodevelopment disorder. The innovation of econometric model, using meta-regression to fit non-linear relationship between age and cognitive function, could help us to model the curve of cognition change across different age groups, not only in social cognition but also in other domains cognitive domains in future research.

Chapter Four compared the social cognition measured by RMET scale in three groups of participants, drugnaïve patients with schizophrenia, matched treated control patients, and matched healthy controls. It is worthy mentioning that to make RMET scale applicable to the drug-native respondents in this study (the mean DUP over 20 years, mean age over 50, and median of years of schooling only 3 years), we revised the original RMET scale and named it RMET-CV-R.

The successful completion rates were highest in healthy controls, intermediate in treated controls, and lowest in the untreated drug-naïve patients (adjusted completion rates, 95.7%, 69.5, and 50.4% respectively), reflecting that being able to complete the test itself required certain level of social cognitive function. Due to the unbalanced complete status, simply using linear regression would introduce selection bias and overestimate the RMET score in patients, so Heckman estimation was used in this study to control for the non-random missing. The adjusted RMET-CV-R scores using Heckman estimation in three groups was 44.5% in UT group, 31.3% in TC group, and 23.1% in HC group, and the difference between any two groups was significant.

Though both of the patients groups had lower RMET score than healthy control indicating the social cognitive impairment in patients with schizophrenia, patients with regular treatment still show much better social cognition function. Good adherence of antipsychotic helps patients with schizophrenia to preserve some level of social cognition.

This study unprecedently evaluated the social cognition of chronic drug-naïve patients with schizophrenia who we had very little information about. This study not only confirmed that anti-psychotic medication adherence reduced the impairment of social cognition in patients with schizophrenia, it was also the first study quantifying the social cognitive deficits of those patients.

However, this study suffers from its cross-section nature. There are quite a number of questions requires to collect more data, such as longitudinal data to follow up these patients, to be answered: 1) Is anti-psychotic medication still effective for these patients who had a very long duration of untreated psychosis; 2) Does the cognition of these drug-native patients decline faster; 3) How to evaluate the cognitive function of patients who are not able to complete the cognitive scale.

In conclusion, the mental health gap of schizophrenia in rural area has been alleviated by implementing community mental health services and regularly using anti-psychotic medication is still an effective way to for patients to control symptoms and improve cognitive function. However, there are still a large number of patients with schizophrenia not regularly using medication due to lack of mental health literacy or access to medication, and they are suffered from more severe symptoms and cognitive impairments.